

UNIVERSITE DE LIMOGES

ECOLE DOCTORALE Science – Technologie – Santé ED258

Faculté des Sciences et Techniques

Institut des Sciences de la Vie et de la Santé

Unité de Génétique Moléculaire Animale

UMR 1061 INRA/Université de Limoges

Thèse N° [47-2006]

Thèse

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE DE LIMOGES

Discipline : Biologie, Sciences, Santé

Présentée et soutenue publiquement par

Céline LORIOL

Le 22 septembre 2006

Les *O*-fucosyltransférases

Caractérisation des enzymes bovines et étude préliminaire du rôle de Pofut1 murine dans la différenciation de la cellule musculaire.

Directeurs de Thèse : Pr Abderrahman Maftah & Dr Agnès Germot

Jury :

Rapporteurs : **Jean-Claude Michalski**, Directeur de Recherche, INSERM, Villeneuve d'Ascq
Rafael Oriol, Directeur de Recherche, CNRS, Villejuif

Examineurs : **Isabelle Martelly**, Professeure, Université Paris XII, Créteil
Raymond Julien, Professeur, Université de Limoges
Agnès Germot, Maître de Conférences, Université de Limoges
Abderrahman Maftah, Professeur, Université de Limoges

UNIVERSITE DE LIMOGES

ECOLE DOCTORALE Science – Technologie – Santé ED258

Faculté des Sciences et Techniques

Institut des Sciences de la Vie et de la Santé

Unité de Génétique Moléculaire Animale

UMR 1061 INRA/Université de Limoges

Thèse N° [47-2006]

Thèse

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE DE LIMOGES

Discipline : Biologie, Sciences, Santé

Présentée et soutenue publiquement par

Céline LORIOL

Le 22 septembre 2006

Les *O*-fucosyltransférases

Caractérisation des enzymes bovines et étude préliminaire du rôle de Pofut1 murine dans la différenciation de la cellule musculaire.

Directeurs de Thèse : Pr Abderrahman Maftah & Dr Agnès Germot

Jury :

Rapporteurs : **Jean-Claude Michalski**, Directeur de Recherche, INSERM, Villeneuve d'Ascq
Rafael Oriol, Directeur de Recherche, CNRS, Villejuif

Examineurs : **Isabelle Martelly**, Professeure, Université Paris XII, Créteil
Raymond Julien, Professeur, Université de Limoges
Agnès Germot, Maître de Conférences, Université de Limoges
Abderrahman Maftah, Professeur, Université de Limoges
Abderrahman Maftah, Professeur, Université de Limoges

*A mes parents,
A Lionel.*

Remerciements

Après quatre années de travail, il est difficile de trouver les mots justes pour remercier tous ceux qui y ont contribué et m'ont apporté leur soutien, mais je vais essayer...

Je tiens tout d'abord à remercier le professeur Abderrahaman Maftah, directeur de cette thèse, pour m'avoir accueilli, il y a maintenant cinq ans au sein de l'EA3176 ensuite regroupée avec l'URM-INRA 1061. Merci Abdou de m'avoir toujours accordé votre confiance dans le travail. Merci de m'avoir accompagné au cours de ces années et d'avoir toujours été disponible pour moi, pour quoi que ce soit. Merci pour votre soutien, vos encouragements et surtout votre enthousiasme et votre optimisme. J'ai beaucoup de chance de vous avoir eu comme directeur de thèse.

J'adresse tous mes remerciements à Agnès Germot, Maître de conférences et co-directrice de cette thèse. Agnès, comment trouver les mots justes pour te remercier, c'est impossible, je crois. Tu le sais, les choses qui ne sont pas dites sont souvent les plus importantes... Grâce à toi j'ai découvert les joies de l'enseignement, j'ai appris à être critique sur mon travail et à toujours aller au bout des choses. Tu m'as fait découvrir les joies de la phylogénie, un domaine aussi passionnant que complexe. Merci pour ta patience durant les moments difficiles et pour tes nombreux conseils et le travail que nous avons effectué ensemble. Tu as su me guider au mieux au cours de toutes ces années. Grâce à toi je suis fière de ma thèse (je sais que tu comprendras). MERCI pour tout, infiniment.

Merci aux membres du jury d'avoir accepté de juger ce travail. M. Jean-Claude Michalsky et M. Rafael Oriol, rapporteurs. Mme Isabelle Martelly, examinatrice. Merci à M. Raymond Julien, également, d'en être le Président.

Je remercie l'ensemble des membres du laboratoire ayant participé de près ou de loin à l'avancement de ma thèse. Merci mon petit Pascal pour ton soutien, ta disponibilité et ton amitié. Merci Marie-pierre pour ton assistance technique, ta disponibilité et ta bonne humeur. Merci Chantal pour ces moments passés à discuter dans la salle du cytomètre. Merci François pour ta gentillesse et ton soutien. Merci Fabrice pour ton aide précieuse de tous les jours, pour tes encouragements, ton soutien, pour nos discussions, souvent acharnées...Merci aux végétalistes, Gaëlle et Catherine ou encore « boulet et bourgeon » pour toutes nos discussions et rigolades lors de pauses dans la salle fumeur de plein air. Merci à tous les autres statutaires et étudiants du laboratoire pour leur bonne humeur. J'adresse une mention particulière à Aymeric, merci pour ton aide tout au long de cette année de thèse, merci également à Mathilde. A tous bonne continuation et bon courage pour la suite. Merci Florence pour ton amitié et ta joie de vivre. Merci à Seb, Christelle, Barbara et Claire pour leur amitié, leur écoute et leurs encouragements.

Merci aux anciens du labo. Ah la fine équipe !!
Merci les filles pour toutes nos soirées hebdomadaires de l'an dernier, soirées riches en discussions et surtout en potins, n'est-ce pas loulou ? Merci Mélanie et Sandrine pour votre

gentillesse. Merci Carole pour ton accueil chaleureux au sein du laboratoire dès le premier jour, tu es formidable, merci pour nos discussions, tes conseils, ton soutien et ton amitié. Merci Lydie pour ta joie de vivre et ta bonne humeur, tu feras une super instit. Ma petite Peggy, que te dire de plus que tu ne saches déjà ! Je t'adore, merci pour nos soirées pyjama toutes les deux dans mon appart à discuter en regardant un film avec qui tu sais... Ne changes pas.

Merci Sylvain pour ton amitié et tes conseils. Merci Mika pour nos discussions, nos sorties tous les quatre et nos chorégraphies sur de la musique electro. Merci Christophe et Vincent pour vos jeux de mots à 2 euros.

Merci à mes parents de tout ce qu'ils ont fait pour moi. Tout l'amour et le bonheur que vous me donnez m'ont porté pendant ces longues années. J'ai une chance fabuleuse d'avoir des parents comme vous.

Merci Lionel, cela fait maintenant 10 ans que tu me supportes et ce n'est pas fini !!!! Merci pour ta tendresse, ta compréhension, ta patience, ta confiance en moi, ton soutien de tous les jours pendant ces années de thèse et pour tout le reste... Je ne peux évidemment oublier de remercier ta famille. Merci Dominique et William, vous êtes des beaux parents super. Merci à ta grand mère, tous tes oncles, tantes, cousins et cousines pour leur formidable accueil dans la famille il y a longtemps maintenant. Merci pour tous ces moments avec vous.

Merci à tous ceux qui étaient présents au barbecue du labo. N'ayez aucune crainte, au pot de thèse, je promets de ne pas courir avec une bouteille en verre à la main. Il n'est pas facile de terminer sa rédaction avec un seul doigt ! Laure, une vocation ratée, tu ferais une super infirmière. Merci Arnaud quand tu auras terminé tes études tu seras mon médecin traitant. Merci de vous être tous si bien occupés de moi ce soir là.

Je terminerai en remerciant tous ceux qui, bien que non cités, conserveront une place chère à mon cœur.

Merci à tous et à toutes.

Publications et communications

Publications

Loriol, C., Dupuy, F., Rampal, R., Dlugosz, M.A., Haltiwanger, R.S., Maftah, A., and Germot, A. (2006) Molecular evolution of protein *O*-fucosyltransferase genes and splice variants. *Glycobiology* 16, 736-747.

Loriol, C., Audfray, A., Dupuy, F., Germot, A., and Maftah, A. (2007) The two *N*-glycans present on bovine Pofut1 are differently involved in its solubility and activity. *FEBS Journal* 274, 1202-1211.

Communications écrites

Loriol, C., Dupuy, F., Rampal, R., Haltiwanger, R.S., Maftah, A., and Germot, A. Tracking the link between *O*-fucosylation and differentiation of skeletal muscle cells. 7th Jenner Glycobiology and Medicine Symposium, 5-8 Septembre 2004, Oxford (Angleterre).

Loriol, C., Dupuy, F., Rampal, R., Haltiwanger, R.S., Germot, A., and Maftah, A. Does *O*-fucosylation regulate differentiation of skeletal muscle cells ? 4th International Symposium on Glycosyltransferases-GlycoT, 4-7 Novembre 2004, Le Touquet (France).

Communications orales

Loriol, C. *Rôle de la O-fucosylation dans le réseau d'interactions moléculaires contrôlant la myogenèse.* Séminaire des Thésards du département de Génétique Moléculaire Animale de l'INRA, Mai 2003, Jouy-en-Josas (France).

Loriol, C. *O-fucosylation et Myogenèse. Etude des gènes pofut chez Bos taurus.* Réunion du Groupement de Recherche CNRS-2590 : Génomique et Génie des Glycosyltransférases, Novembre 2003, Université de Rouen (France).

Loriol, C. *Rôle de la O-fucosylation pendant la Myogenèse.* Séminaire des Thésards du département de Génétique Moléculaire Animale de l'INRA, Mai 2004, Tours (France).

Loriol, C., Dupuy, F., Rampal, R., Haltiwanger, R.S., Maftah, A., and Germot, A. Tracking the link between *O*-fucosylation and differentiation of skeletal muscle cells. 7th Jenner Glycobiology and Medicine Symposium, 5-8 Septembre 2004, Oxford (Angleterre).

Loriol, C., Dupuy, F., Rampal, R., Haltiwanger, R.S., Germot, A., and Maftah, A. *Does O-Fucosylation regulate differentiation of skeletal muscle cells ?* 4th International Symposium on Glycosyltransferases-GlycoT, 4-7 Novembre 2004, Le Touquet (France).

Dupuy, F., Forestier, L., Loriol, C., Charreau, S., Germot, A. et Maftah, A. Etude du rôle de la glycosylation dans le processus de myogenèse. Réunion du Groupement de Recherche CNRS-2590 : Génomique et Génie des Glycosyltransférases, Avril 2005, Orsay (France).

Liste des Abréviations

aa : acide aminé	dNTP : Désoxynucléotide triphosphate
ADN : Acide désoxyribonucléique	DSL : Protéines transmembranaires <u>D</u> elta, <u>S</u> errate, <u>L</u> ag2
ADNc : ADN complémentaire	ECN : Extracellular Notch
ADP : Adénosine diphosphate	EDL : Extensor Digitorum Longus
ANK : Domaines ankyrine répétés	EDTA : Ethylenediaminetetraacetic acid
ARN : Acide ribonucléique	EGF : Epidermal Growth Factor repeat
ARNm : ARN messenger	eGFP/IGFP : enhancer/long-lasting Green Fluorescent Protein
ARNnc : ARN non codant	ES : Embryonic stem (cell)
bHLH : basic Helix-Loop-Helix	EST : Expressed Sequence Tag
BiP : Binding protein	Ex : Embryonic day x, (x = days post coïtum)
BMP : Bone Morphogenetic Protein	FACS : Fluorescence Activated Cell Sorter
C2C12 : Lignée myoblastique murine	FCMD : Fukuyama Congenital Muscular Dystrophy
CBF1-RBP : C Promoter Binding Factor 1-Retinol-Binding Protein	FCS : Fœtal Calf Serum
CDG : Congenital Disorder of Glycosylation	FGF : Fibroblast Growth Factor
CHO : Lignée cellulaire dérivée de cellules ovariennes de hamster chinois	FITC : Fluorescein isothiocyanate
CMD1C, 1D : Congenital Muscular Dystrophy type 1C, 1D	FKRP : Fukutin-related protein
CMV : Cytomégalovirus	Fuc : Fucose
CNX : Calnexine	GAG : Glycosaminoglycane
COS : Lignée cellulaire dérivée de cellules de rein de singe vert d'Afrique, transformées par l'antigène T du SV40	Gal : Galactose
CRT : Calréticuline	GalNac : <i>N</i> -actétylgalactosamine
Ct : Threshold Cycle	GalNac-T : <i>N</i> -actétylgalactosaminyltransférase
Da, KDa : Dalton, KiloDalton	Gal-T : Galactosyltransférase
DGC : Dystrophin-associated Glycoprotein Complex	GDP : Guanosine diphosphate
DII : Delta-like	Glc : Glucose
DMEM : Dulbecco's Modified Eagle Medium	GlcA : Acide glucuronique
	GlcNAc : <i>N</i> -Acétylglucosamine
	GlcNAc-T : <i>N</i> -Acétylglucosaminyltranférase
	GlcT : Glucosyltransférase

GM130 : Golgi matrix protein 130	ORF : Open Reading Frame
GPI : Glycosyl Phosphatidyl Inositol	OST : Oligosaccharyltransférase
HEPES : N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)	p/v : poids / volume
HES : Hairy/Enhancer of Split	PAGE : Polyacrylamide Gel Electrophoresis
HGF : Hepatocyte Growth Factor	Pax : Paired box
ICN : Intracellular Notch	pb, Kpb : paire de bases, Kilo paires de bases
IGF : Insulin-like Growth Factor	PBS : Phosphate Buffer Saline
IgG : Immunoglobuline G	PCR : Polymerase Chain Reaction
KO : Knock Out	PE : Phycoérythrine
L6 : Lignée myoblastique de rat	PNGase : Peptide-N-Glycosidase F
LADII : Leukocyte Adhesion Deficiency type II	PSA-NCAM : Polysialic Acid-Neural Cell Adhesion Molecule
Lbx1 : Ladybird homeobox homolog 1	RAM : Regulation of Amino acid Metabolism
Le : Antigène Lewis	RE : Réticulum endoplasmique
Lfng : Lunatic Fringe	RER : Réticulum endoplasmique rugueux
LGMD2I : Limb Girdle Muscular Dystrophy type 2I	Rfng : Radical Fringe
LNR : Lin12-Notch repeat	RNAi : Ribonucleic acid interference
MADS-box : MCM1-Agamous-Deficiens-SRF box	RT-PCR : Reverse Transcription-PCR
Man : Mannose	SDS : Sodium Dodecyl Sulfate
MB, GB : Mégabase, Gigabase	Shh : Sonic Hedgehog
MEB : Muscle-Eye-Brain disease	shRNA : Short hairpin ribonucleic acid
MEC : Matrice extracellulaire	sialyl-Tn : NeuAc α 1,6GalNAc α -S/T
MEF2 : Myocyte Enhancer Factor 2	SiaT : Sialyltransférase
Mfng : Manic Fringe	Six : Sine oculis homeobox
Mox2 : Mesenchyme homeobox 2	sLe : Antigène sialyl-Lewis
MRF4 : Muscle Regulatory Factor 4 (synonyme de Myf6)	SRF : Serum Response Factor
Msx1 : Muscle segmentation homeobox 1-like protein	Su(H) : Suppressor of hairless
myd : myodystrophy	SV40 : Simian Virus 40
Myf : Myogenic Factor	TAD : Transcription Activation Domain
MyoD : Myoblast Determination gene	TGFβ : Transforming Growth Factor β
NCR : Notch Cytokin Response	TLDA : TaqMan Low Density micro-Array
NeuAc : N-acetyl-D-neuraminic acid	Tn : GalNAc α -S/T
NLS : Nuclear Localisation Signal	t-PA : tissue Plasminogen Activator
NTM : Notch Transmembrane Fragment	TSP1 : Thrombospondine 1
OGT : O-GlcNAc Transférase	TSR : Thrombospondin type 1 repeat
	UDP : Uridine diphosphate

Abréviations

UMP : Uridine monophosphate

UTR : Untranslated Region

v/v : volume / volume

VWC : von Willebrand factor type C domain

WWS : Walker-Warburg Syndrom

Wnt : Wingless-type MMTV integration site family
member

Xyl : Xylose

Résumé

La *O*-fucosylation est l'ajout d'un fucose sur un résidu sérine ou thréonine compris dans deux types de domaines peptidiques, les EGF et les TSR. Cette modification post-traductionnelle dépend de deux enzymes, Pofut1 pour les EGF et Pofut2 pour les TSR. Nous avons étudié l'évolution des gènes *Pofut1* et *Pofut2*, et avons démontré que ces gènes présents en un seul exemplaire, existaient déjà chez l'ancêtre des Bilatériens voire des Métazoaires, probablement sous forme morcelée. Une situation originale existe pour *Pofut2*, retrouvé chez un groupe de Protozoaires, les Apicomplexés. La structure des deux gènes chez le bovin et leur expression tissulaire ont été établies. Nous trouvons une situation inédite puisqu'il existe pour chacun d'entre eux, cinq variants transcriptionnels dont un seul code l'enzyme active, différemment exprimés selon les tissus bovins. Les enzymes Pofut1 et Pofut2 actives seraient présentes dans tous les tissus analysés, à l'exception notable des muscles squelettiques de l'adulte où des formes atypiques sont présentes. Les autres variants transcriptionnels, la plupart tronqués, auraient un rôle dans la régulation du taux d'expression des gènes. Pofut1 est la première fucosyltransférase à avoir été identifiée comme résidente du réticulum endoplasmique. D'un point de vue fonctionnel, nous avons démontré chez le bovin que cette glycosyltransférase, porteuse de deux *N*-glycanes vraisemblablement de type oligomannosidique, était correctement repliée et donc douée d'une activité enzymatique, à condition que son premier site de *N*-glycosylation soit occupé. L'activité des récepteurs Notch et de leurs ligands présents à la surface de nombreuses cellules est modulée par l'état de *O*-fucosylation de leurs domaines EGF. Etant donné l'implication de ces récepteurs dans la régulation de la myogenèse, la *O*-fucosylation doit y contribuer tout autant. Ainsi, nous avons débuté une étude des répercussions d'une modulation de l'expression de Pofut1 sur la différenciation de cellules primaires myoblastiques bovines et de cellules murines de la lignée C2C12. Nous montrons que la surexpression transitoire de l'enzyme Pofut1 murine retarde l'expression des facteurs de transcription myogéniques de la famille bHLH: Myf5, MyoD, Myogénine et MRF4. Ces résultats, encore préliminaires, ouvrent de nouvelles et passionnantes perspectives d'analyse de l'influence des *O*-fucosyltransférases au cours du processus de myogenèse.

Mots clés : *O*-fucosylation, organisation génomique, variants transcriptionnels, phylogénie, activité *O*-fucosyltransférase, myogenèse, bovin

Abstract

O-fucosylation is the addition of a fucose on serine or threonine comprised in two types of peptidic domains, EGF and TSR. This post-translational modification depends on two enzymes. Pofut1 is responsible for *O*-fucosylation of EGF repeats, whereas Pofut2 adds fucose on TSR. We studied the evolution of *Pofut1* and *Pofut2* genes and demonstrated that these genes, present in a single copy, already existed in the ancestor of Bilaterians, or even Metazoans, probably in polyexonic form. An original situation exists for *Pofut2*, recovered in a group of Protozoans, the *Apicomplexa*. Structures of the two bovine genes and their tissular expression have been established. We find an original situation since it exists for each of them, five transcript variants of which only one encodes the active enzyme, differently expressed among bovine tissues. The Pofut1 and Pofut2 active enzymes would be present in all analyzed tissues, except for adult skeletal muscles where atypical forms are present. The other transcript variants, more or less truncated, probably play a role in regulating the expression level of these genes. Pofut1 is the first fucosyltransferase to have been identified as an endoplasmic reticulum resident. Functionally, we demonstrated for the bovine that this glycosyltransferase, bearing two *N*-glycans probably of oligomannosidic type, was correctly folded and therefore possessed an enzymatic activity, provided its first site of *N*-glycosylation is occupied. The activity of the Notch receptors and their ligands, present at the surface of numerous cells, is modulated by the state of *O*-fucosylation of their EGF domains. Considering the implication of these receptors in the regulation of myogenesis, *O*-fucosylation must contribute all as much there. Thus, we started a study of the repercussions of a modulation of Pofut1 expression on the differentiation of bovine primary muscular cells and murine cells of the C2C12 lineage. We show that the transient surexpression of the murine Pofut1 enzyme delays the expression of myogenic transcription factors belonging to bHLH family: Myf5, MyoD, Myogenine and MRF4. These preliminary results open new and exciting perspectives of analysing the influence of *O*-fucosyltransferases during the complex process of myogenesis.

Key words : *O*-fucosylation, genomic organisation, transcript variants, phylogeny, *O*-fucosyltransferase activity, myogenesis, cattle

Sommaire

AVANT-PROPOS	1
EXPOSE BIBLIOGRAPHIQUE	9
1 UNE MODIFICATION POST-TRADUCTIONNELLE MAJEURE DES PROTEINES, LA GLYCOSYLATION	11
1.1 PRESENTATION ET FONCTIONS DE LA <i>N</i> -GLYCOSYLATION DES PROTEINES	15
1.1.1 <i>Structures et diversité des N-glycanes</i>	15
1.1.2 <i>Biosynthèse des N-glycanes</i>	18
1.1.2.1 Dans le réticulum endoplasmique rugueux	19
1.1.2.2 Dans l'appareil de Golgi	22
1.1.3 <i>Rôles biologiques des N-glycanes</i>	22
1.1.3.1 Vis-à-vis de la <i>N</i> -glycoprotéine	22
1.1.3.2 Vis-à-vis de l'organisme	25
1.2 PRESENTATION ET FONCTIONS DES DIFFERENTS TYPES DE <i>O</i> -GLYCOSYLATION DES PROTEINES	27
1.2.1 <i>La O-glycosylation de type GAG</i>	28
1.2.2 <i>La O-glycosylation de type mucine</i>	30
1.2.3 <i>La O-mannosylation</i>	32
1.2.4 <i>La O-N-acétylglucosaminylation</i>	34
1.2.5 <i>La O-glucosylation</i>	34
1.2.6 <i>La O-fucosylation</i>	36
1.2.7 <i>Les différents types de O-glycanes ont des rôles biologiques communs</i>	36
1.2.7.1 Vis-à-vis de la <i>O</i> -glycoprotéine	36
1.2.7.2 Vis-à-vis de l'organisme	37
2 NOS GLYCOSYLTRANSFERASES D'INTERET, LES O-FUCOSYLTRANSFERASES	41
2.1 UNE GLYCOSYLTRANSFERASE IMPLIQUEE DANS LE DEVELOPPEMENT EMBRYONNAIRE, LA <i>O</i> -FUCOSYLTRANSFERASE 1	45
2.1.1 <i>Cibles de l'enzyme Pofut1, les EGF</i>	45
2.1.1.1 Protéines possédant des EGF <i>O</i> -fucosylés	45
2.1.1.2 Formation du <i>O</i> -fucosylglycane porté par les EGF	46
2.1.2 <i>Caractérisation de l'enzyme Pofut1</i>	48
2.1.2.1 Identification de l'activité <i>O</i> -fucosyltransférase 1	48
2.1.2.2 Purification de l'enzyme	48
2.1.2.3 Identification du gène correspondant, <i>Pofut1</i>	49
2.1.2.4 Localisation sub-cellulaire de l'enzyme <i>Pofut1</i>	52

2.1.3	<i>Rôles du O-fucose des domaines EGF</i>	55
2.2	DECOUVERTE RECENTE D'UNE ENZYME HOMOLOGUE A POFUT1, LA O-FUCOSYLTRANSFERASE 2	63
2.2.1	<i>Une seconde cible de la O-fucosylation, les TSR</i>	63
2.2.1.1	Protéines possédant des TSR O-fucosylés	63
2.2.1.2	Formation du O-fucosylglycane porté par les TSR	64
2.2.2	<i>Caractérisation de l'enzyme Pofut2</i>	65
2.2.2.1	Identification de l'activité O-fucosyltransférase 2	65
2.2.2.2	Etablissement de la structure du gène <i>Pofut2</i>	66
2.2.2.3	Localisation sub-cellulaire de l'enzyme Pofut2	68
2.2.3	<i>Rôles du O-fucose des domaines TSR</i>	68
2.3	CARACTERISTIQUES MOLECULAIRES ET FONCTIONNELLES DES O-FUCOSYLTRANSFERASES	71
2.4	EXISTE T'IL UNE TROISIEME O-FUCOSYLTRANSFERASE ?	75
2.5	O-FUCOSYLATION ET DEVELOPPEMENT EMBRYONNAIRE	77
	POURQUOI GLYCOSYLATION ET MYOGENESE ?	81
3	LA MYOGENESE DU MUSCLE SQUELETTIQUE	89
3.1	ORIGINE EMBRYONNAIRE DU MUSCLE SQUELETTIQUE	93
3.2	DETERMINATION MYOGENIQUE	97
3.2.1	<i>Détermination et migration des prémyoblastes</i>	97
3.2.2	<i>Différenciation des précurseurs musculaires</i>	99
3.2.2.1	Les protéines à domaine bHLH	99
3.2.2.1.1	Facteurs myogéniques (MRFs)	99
3.2.2.1.2	Cinétiques d'expression	100
3.2.2.1.3	Activités	101
3.2.2.2	Les protéines à domaine MADS	104
3.2.2.2.1	Facteur de réponse au sérum (SRF)	104
3.2.2.2.2	Famille des Myocyte Enhancer Factors 2 (MEF2)	105
3.3	EFFECTEURS DE LA MYOGENESE	107
3.3.1	<i>Les inducteurs</i>	107
3.3.1.1	La famille des Insulin-like growth factors (IGFs)	107
3.3.1.2	La famille Wntless (WNTs)	108
3.3.2	<i>Les inhibiteurs</i>	109
3.3.2.1	La famille des Transforming growth factors β (TGF β)	109
3.3.2.2	La famille des Fibroblast growth factors (FGFs)	110
3.3.2.3	Notch et ses ligands	111
3.4	LA MYOGENESE <i>IN VITRO</i>	115
	RESULTATS-DISCUSSION	117
1	EVOLUTION DES GENES DE O-FUCOSYLTRANSFERASES. STRUCTURE ET EXPRESSION	
	DES GENES POFUT1 ET POFUT2 CHEZ BOS TAURUS.	121
1.1	PUBLICATION 1	125
1.2	RESULTATS COMPLEMENTAIRES	129
1.2.1	<i>Expression tissulaire des variants transcriptionnels</i>	129

1.2.2	<i>Production d'anticorps anti-Pofut bovins</i>	134
1.2.3	<i>Activités des protéines codées par les différents variants</i>	138
1.2.4	<i>Evolution des gènes de O-fucosyltransférases</i>	143
1.2.5	<i>Synthèse et conclusion</i>	152
2	CONSERVATION DES SITES DE N-GLYCOSYLATION DE POFUT1, RELATION	
	STRUCTURE/FONCTION	169
2.1	PUBLICATION 2	173
2.2	RESULTATS COMPLEMENTAIRES	175
2.2.1	<i>Expression tissulaire de l'enzyme bovine Pofut1</i>	175
2.2.2	<i>Expression tissulaire de l'enzyme bovine Pofut2</i>	178
2.2.3	<i>Synthèse et conclusion</i>	181
3	POFUT1 : UN ROLE DANS LA DIFFERENCIATION DE CELLULES MUSCULAIRES EN	
	CULTURE ?	183
3.1	MISE EN PLACE DE LA CULTURE PRIMAIRE DE CELLULES SATELLITES BOVINES ET DES VECTEURS D'EXPRESSION EUCARYOTE	187
3.1.1	<i>Isolement, culture et différenciation</i>	187
3.1.2	<i>Construction des vecteurs eucaryotes pour les transfections</i>	189
3.1.2.1	Vecteur de surexpression	189
3.1.2.2	Vecteur de shRNA	191
3.1.3	<i>Transfection des cellules bovines en culture primaire</i>	194
3.2	ETUDE DU ROLE DE POFUT1 DANS LA DIFFERENCIATION DE CELLULES C2C12	199
3.2.1	<i>Construction du vecteur de surexpression</i>	199
3.2.2	<i>Culture et transfection</i>	200
3.2.2.1	Difficultés rencontrées	200
3.2.2.2	Utilisation d'autres vecteurs	203
3.2.3	<i>Premiers résultats obtenus</i>	205
3.2.3.1	Surexpression de Pofut1 dans les cellules C2C12	205
3.2.3.1.1	Transfection et induction de la différenciation	205
3.2.3.1.2	Vérification de l'expression de Pofut1	205
3.2.3.1.3	Répercussion de la surexpression de Pofut1 sur l'expression des facteurs myogéniques bHLH	210
3.2.3.2	Inhibition de l'expression de Pofut1	214
3.3	MATERIELS ET METHODES	215
	Réaction de PCR	217
	Méthodes relatives aux cellules	217
	Dosage des ARN	220
	RT-PCR semi-quantitative en temps réel	220
	CONCLUSION ET PERSPECTIVES	223
	NOUVEAU CHAPITRE DE LA THESE	229
	REFERENCES BIBLIOGRAPHIQUES	243

Index des figures

Figure 1 : Exemples de structure des trois classes principales de <i>N</i> -glycanes présentes chez les Eucaryotes.....	16
Figure 2 : Différentes ramifications trouvées sur les <i>N</i> -glycanes complexes.	18
Figure 3 : Voie de biosynthèse du précurseur oligosaccharidique des <i>N</i> -glycanes dans le RER.	20
Figure 4 : Maturation des <i>N</i> -glycanes dans le réticulum endoplasmique rugueux et l'appareil de Golgi.	21
Figure 5 : Repliement d'une protéine glycosylée dans le RER. Cycle de déglucosylation/ reglucosylation.	23
Figure 6 : Représentation schématique des principaux monosaccharides <i>O</i> -liés sur des résidus Sérine et Thréonine.	27
Figure 7 : Représentation schématique d'un protéoglycane.	29
Figure 8 : Représentation des quatre principaux noyaux des <i>O</i> -glycanes de type mucine.	30
Figure 9 : Représentation schématique du <i>O</i> -mannosylglycane présent sur l' α -dystroglycane.	33
Figure 10 : Représentation des domaines EGF et TSR.	43
Figure 11 : Structure d'un <i>O</i> -fucosylglycane greffé sur un domaine EGF.	46
Figure 12 : Représentation de l'organisation génomique du gène <i>POFUT1</i> humain et de deux de ses variants transcriptionnels.	50
Figure 13 : Représentation de l'organisation génomique du gène <i>Pofut1</i> murin et du variant transcriptionnel décrit.	51
Figure 14 : Représentation schématique d'une glycosyltransférase transmembranaire de type II de l'appareil de Golgi.	52
Figure 15 : Immunolocalisation de la protéine recombinante POFUT1 humaine dans les cellules COS-1.	53
Figure 16 : Représentation schématique de la translocation et de la séquestration de protéines solubles dans le RER.	54
Figure 17 : Modélisation du domaine EGF-like de Cripto murin et du premier domaine EGF du facteur de coagulation VII Humain.	56
Figure 18 : Domaines conservés des récepteurs Notch chez la drosophile (dNotch) et les Mammifères (mNotch1 à 4).	58
Figure 19 : Domaines conservés des ligands de Notch chez la drosophile et les Mammifères.	59
Figure 20 : Structure d'un <i>O</i> -fucosylglycane greffé sur un domaine TSR.	65
Figure 21 : Représentation de l'organisation génomique du gène <i>POFUT2</i> humain et de ses variants transcriptionnels.	66
Figure 22 : Représentation de l'organisation génomique du gène <i>Pofut2</i> murin et de son variant transcriptionnel.	67
Figure 23 : Biosynthèse du GDP-fucose.	72
Figure 24 : Représentation schématique des enzymes POFUT1 et POFUT2 humaines.	73
Figure 25 : Alignement des motifs conservés d' α 2, α 6- et <i>O</i> -fucosyltransférases.	73

Figure 26 : Expression des gènes impliqués dans les voies de signalisation par Notch chez les embryons de souris sauvages (gauche) et mutants <i>Pofut1</i> <i>-/-</i> (droite).	79
Figure 27 : Représentation schématique du complexe DGC.	85
Figure 28 : Représentation schématique de la somitogenèse chez les Vertébrés.....	93
Figure 29 : Représentation schématique de la somitogenèse chez les Vertébrés.....	94
Figure 30 : Schéma illustrant la séquence des évènements qui aboutissent à la formation des fibres musculaires striées.	96
Figure 31 : Représentation schématique de la formation du muscle squelettique au niveau des membres chez les Vertébrés.	97
Figure 32 : Représentation schématique du facteur de transcription MyoD.	100
Figure 33 : Représentation schématique de l'expression des gènes myogéniques de la famille MyoD, au cours du développement musculaire chez la souris.	100
Figure 34 : Résultats des invalidations des facteurs myogéniques chez la souris.....	103
Figure 35 : Expression des facteurs myogéniques de la famille MyoD au cours de la myogenèse.	103
Figure 36 : Schéma récapitulatif de l'effet inducteur (vert) ou inhibiteur (rouge) des différents effecteurs de la myogenèse, secrétés par les structures environnant les somites.	111
Figure 37 : Représentation schématique des voies de signalisation du récepteur Notch contrôlant la myogenèse.....	113
Figure 38 : Expression tissulaire des variants transcriptionnels de <i>Pofut2</i> dans les tissus de bovin adulte.	130
Figure 39 : Expression tissulaire des variants transcriptionnels de <i>Pofut1</i> et <i>Pofut2</i> dans les tissus d'embryon bovin. Comparaison avec l'expression tissulaire chez l'adulte.	131
Figure 40 : Représentation schématique de l'organisation génomique du gène <i>POFUT2</i> humain et de l'épissage alternatif conduisant aux transcrits référencés dans les bases de données moléculaires (variants A, B et C, et ESTs).	133
Figure 41 : Analyse par SDS-PAGE 10,5% de l'expression des enzymes recombinantes bovines <i>Pofut1A</i> et <i>Pofut2B</i>	135
Figure 42 : Représentation schématique des protéines <i>Pofut1A</i> et <i>Pofut2B</i> bovines produites dans le système bactérien.	135
Figure 43 : Vérification de la spécificité des anticorps dirigés contre les enzymes bovines <i>Pofut1</i> et <i>Pofut2</i>	136
Figure 44 : Représentation schématique des protéines bovines <i>Pofut1A</i> , <i>Pofut2A</i> et <i>Pofut2B</i> produites dans les cellules COS-1.....	138
Figure 45 : Alignement des séquences protéiques déduites des cinq variants transcriptionnels du gène bovin <i>Pofut1</i>	139
Figure 46 : Alignement des séquences protéiques déduites des cinq variants transcriptionnels du gène bovin <i>Pofut2</i>	140
Figure 47 : Activité <i>O</i> -fucosyltransférase 1 des protéines codées par les variants bovins <i>Pofut1a</i> à <i>Pofut1e</i>	141
Figure 48 : Activité <i>O</i> -fucosyltransférase 2 des protéines codées par les variants bovins <i>Pofut2a</i> à <i>Pofut2e</i>	142
Figure 49 : Organisations exon/intron des CDS des gènes <i>Pofut1</i> chez différentes espèces animales.....	145
Figure 50 : Organisations exon/intron des CDS des gènes <i>Pofut2</i> chez différentes espèces animales.....	146
Figure 51 : Représentation schématique des différentes hypothèses permettant d'expliquer la présence et l'évolution du gène <i>Pofut2</i> chez trois genres de protozoaires du phylum des Apicomplexés.....	149
Figure 52 : Profil d'expression tissulaire de la protéine bovine <i>Pofut1</i>	175

Figure 53 : Expression de Pofut1 dans des tissus de Mammifères.	178
Figure 54 : Profil d'expression tissulaire de la protéine bovine Pofut2.	179
Figure 55 : Expression de Pofut2 dans deux tissus murins.	180
Figure 56 : Résultat d'un tri sur un ensemble de cellules isolées d'un muscle squelettique d'embryon bovin.	188
Figure 57 : Photographies en lumière blanche et immunomarquages des myoblastes et myotubes bovins.	189
Figure 58 : Vecteur choisi pour la surexpression de l'enzyme Pofut1A dans les cellules primaires de muscle bovin.	190
Figure 59 : Principe du shRNA.	192
Figure 60 : Vecteur pSilencer 2.1-U6 de shRNA.	193
Figure 61 : Vecteur pSuper.gfp de shRNA.	194
Figure 62 : Cytogrammes biparamétriques (rapport taille vs fluorescence) obtenus lors d'une expérience de tri cellulaire après transfection de culture primaire de myoblastes bovins.	195
Figure 63 : Western blot anti-Pofut1 réalisé sur 50µg d'extraits protéiques de cellules primaires transfectées par le plasmide pVITRO2-IGFP- <i>Pofut1a</i> bovin et triées.	196
Figure 64 : Photographies en lumière blanche de cellules myoblastiques bovines en phase proliférative.	197
Figure 65 : Cytogrammes biparamétriques (rapport taille vs fluorescence) obtenus lors d'une expérience de tri cellulaire après transfection des cellules C2C12 avec le plasmide pVITRO2-IGFP/ <i>Pofut1</i> murin.	201
Figure 66 : Vecteur témoin pmaxGFP fourni avec le kit de transfection Amaxa.	202
Figure 67 : Cytogramme biparamétrique (rapport taille vs fluorescence) obtenus lors d'une expérience de tri cellulaire après transfection des cellules C2C12 avec le plasmide pmaxGFP.	202
Figure 68 : Principe de la PCR semi-quantitative en temps réel avec une sonde TaqMan... ..	204
Figure 69 : Analyse en RT-PCR semi-quantitative des transcrits de <i>Pofut1</i> dans les cellules C2C12 non transfectées (marron) et dans les cellules transfectées avec la construction pcDNA/TOPO3.1- <i>Pofut1</i> murin (jaune).	206
Figure 70 : Western blot réalisé avec l'anticorps anti-Pofut1 sur 50µg d'extraits de cellules C2C12, 24 heures après transfection avec le plasmide pcDNA/TOPO3.1- <i>Pofut1</i> murin.	207
Figure 71 : Photographies, au microscope inversé (10X) en lumière blanche, des cellules C2C12 non transfectées, au cours de la cinétique de différenciation.	208
Figure 72 : Photographies, au microscope inversé (10X) en lumière blanche, des cellules C2C12 transfectées par le vecteur pcDNA/TOPO3.1- <i>Pofut1</i> murin, au cours de la cinétique de différenciation.	209
Figure 73 : Analyse en RT-PCR semi-quantitative du niveau d'expression du marqueur de prolifération CD34 et des quatre principaux facteurs bHLH intervenant au cours de la myogenèse, Myf5, MyoD, Myogénine et MFR4.	211
Figure 74 : Représentation d'une amplification par PCR en temps réel, d'un fragment d'ADN sur une échelle logarithmique.	221

Index des tableaux

Tableau 1 : Constantes cinétiques de l'enzyme Pofut1 purifiée à partir de cellules CHO et de l'enzyme Pofut1 humaine.....	49
Tableau 2 : Protéines impliquées dans la détermination et la migration des précurseurs musculaires.....	98
Tableau 3 : Couples d'amorces utilisés pour les clonages en vecteur d'expression procaryote pET28c. Les sites de restrictions sont soulignés.....	134
Tableau 4 : Séquences de domaines EGF et TSR O-fucosylables présents dans des protéines (non caractérisées) de <i>Cryptosporidium parvum</i>	150
Tableau 5 : Séquences de domaines EGF et TSR O-fucosylables chez <i>Hydra magnipapillata</i>	151

Avant-Propos

La chimie, la biochimie et la biologie des sucres ont connu un essor à la fin du 19^{ème} siècle, lors de l'étude des processus de fermentation du glucose et du sucrose. C'est l'établissement de la structure du D-glucose en 1891 par Emil Fisher qui marque le début de la recherche sur les glucides.

Pendant longtemps, les glucides ont été associés à la notion de substances de réserve, de composés énergétiques ou de structures de soutien. En effet, on croyait que seuls les acides nucléiques et les protéines portaient des informations biologiques spécifiques. Les glycoconjugués qui résultent de l'association covalente d'une chaîne glucidique nommée glycane, avec une protéine (glycoprotéine) ou avec un lipide (glycolipide - dont nous ne ferons pas état dans ce manuscrit) ont suscité un intérêt croissant en biologie. Ainsi, aux cours des décennies, les chercheurs ont mis en évidence, leur importance dans la détermination des groupes sanguins (Kabat, 1956), leurs rôles dans les interactions cellulaires (Roseman, 1970), leurs implications dans la fécondation (Jego et Boisseau, 1980) et la différenciation erythrocytaire (Liu *et al.*, 1981), etc.

Les glycoconjugués ont pu accéder dès 1988 au rang qui leur était dû, grâce à l'avènement d'une nouvelle discipline scientifique baptisée « Glycobiologie » par Rademacher, Parekh et Dwek. La Glycobiologie est donc une discipline récente alliant les fondements de la biochimie et de la biologie moléculaire des glucides. Son but est de décoder le sens de ce troisième alphabet des sucres, qui s'ajoute à celui des protéines et des acides nucléiques. Ainsi, bien que l'activité biologique des glycoprotéines soit généralement portée par la partie protéique, les glycanes détiennent des informations supplémentaires permettant notamment de moduler la fonction biologique des glycoprotéines, déterminer leur durée de vie, et signaler leur adressage aux différents compartiments cellulaires.

Si la biosynthèse des acides nucléiques et des protéines repose sur la reproduction d'une matrice, celle des glycanes dépend du fonctionnement et de la spécificité de tout un répertoire d'enzymes. Les oligosaccharides sont synthétisés dans le réticulum endoplasmique puis modifiés au sein de l'appareil de Golgi sous l'action coordonnée et séquentielle de diverses glycosyltransférases et glycosidases (Kleene et Berger, 1993). Dans les cellules animales, de nombreux travaux soulignent l'importance de la régulation spatio-temporelle de l'expression de ces enzymes dans la définition des épitopes glycaniques. Grâce au séquençage des génomes, nous savons aujourd'hui qu'environ 1% des gènes d'un organisme leur est dédié (Lowe et Marth, 2003).

En 2002, j'ai commencé ma formation doctorale en rejoignant l'équipe Glycogénome de l'Unité de Génétique Moléculaire Animale (UMR-INRA 1061-Université de Limoges) au sein de la Faculté des Sciences et Techniques de Limoges. Depuis 1995, cette équipe s'intéresse, entre autres, à la caractérisation des gènes bovins de glycosyltransférases (Oulmouden *et al.*, 1997; Mercier *et al.*, 1999; Barraud *et al.*, 2000; Javaud *et al.*, 2000; Loriol *et al.*, 2006), à l'étude des relations structure-fonction des enzymes codées par ces gènes (Dupuy *et al.*, 1999; Holmes *et al.*, 2000; Dupuy *et al.*, 2004) et à l'évolution de ces enzymes dans le règne animal (Wierinckx *et al.*, 1999; Saunier *et al.*, 2001; Dupuy *et al.*, 2002; Javaud *et al.*, 2003; Loriol *et al.*, 2006; Petit *et al.*, 2006). Désormais, elle se consacre en grande partie à l'étude des rôles de ces glycosyltransférases dans des processus développementaux tels que la myogenèse et pathologiques tels que les maladies à prions (Barret *et al.*, 2005).

Au sein du département de génétique animale de l'INRA, notre unité porte un intérêt particulier à la recherche des déterminants génétiques conditionnant le développement musculaire. La compréhension des mécanismes qui régulent la myogenèse pourrait s'avérer primordiale pour une future étude de la variabilité génétique déterminant la masse musculaire chez les animaux de rente. Ainsi, nous nous intéressons aux rôles joués par les glycoconjugués dans les étapes de prolifération et de différenciation de la cellule musculaire. Au cours du développement embryonnaire des Vertébrés, certaines cellules somitiques vont se spécialiser en myoblastes puis se différencier en myotubes qui donneront les fibres musculaires. Cette différenciation est, entre autres, initiée et modulée par les glycoprotéines portées ou secrétées par les structures tissulaires entourant les somites. Plusieurs données

bibliographiques récentes soulignent l'importance fonctionnelle des glycoprotéines (Smalheiser *et al.*, 1998) et des protéoglycanes (Papy-Garcia *et al.*, 2002) dans les processus myogéniques et la physiologie du muscle squelettique (Cieniewski-Bernard *et al.*, 2005). Certaines pathologies neuromusculaires humaines sont notamment dues à des mutations dans des gènes codant des glycosyltransférases (De Bernabe *et al.*, 2003).

Notre équipe a choisi deux stratégies complémentaires, menées en parallèle, afin d'appréhender le rôle de la glycosylation dans le processus de myogenèse. La première stratégie, dite « globale », consiste à étudier l'expression des principaux gènes de la glycosylation au cours de la différenciation de cellules musculaires murines de la lignée C2C12. Cette approche repose sur la mise en place et l'utilisation de la technologie TaqMan Low Density micro-Array, qui autorise l'analyse simultanée par PCR en temps réel, des niveaux d'expression de plusieurs centaines de gènes de la glycosylation. Cette analyse permettra de déterminer les principaux gènes qui influencent le déroulement du programme myogénique chez la souris. Il sera ensuite utile de confirmer l'importance de ces gènes dans un modèle de différenciation de cellules musculaires bovines (myoblastes en culture primaire). A terme, les gènes de la glycosylation jugés cruciaux feront l'objet d'une recherche de polymorphismes potentiellement impliqués dans le phénotype d'hypermusculature chez les bovins. Plus largement, cette étude devrait permettre de révéler les gènes de la glycosylation susceptibles d'expliquer des dérèglements physiologiques (musculature plus ou moins importante) ou pathologiques (atrophie musculaire, myopathie, etc.) affectant la prolifération et la différenciation de la cellule musculaire.

La seconde stratégie concerne l'étude de gènes de la glycosylation qualifiés de « candidats » d'après les connaissances décrites dans la littérature. C'est dans ce cadre que s'inscrivent mes travaux de thèse. Ainsi, il est connu que les récepteurs Notch sont des glycoprotéines impliquées dans l'initiation et la modulation des voies de différenciation embryonnaire. Leur rôle est en particulier démontré dans la différenciation *in vitro*, des cellules myoblastiques de la lignée murine C2C12. Or, la cascade de signalisation intracellulaire déclenchée par l'interaction des récepteurs Notch avec leurs ligands dépend de la présence de *O*-fucoses et de l'allongement de ceux-ci par d'autres sucres. Plus précisément, le processus de *O*-fucosylation que nous avons choisi d'étudier désigne le greffage d'un fucose sur un résidu sérine ou thréonine appartenant à une séquence peptidique consensus. L'étude approfondie de ce type de modification post-traductionnelle est relativement récente (début des années 90) et repose sur les activités de recherche de quelques laboratoires

internationaux situés aux Etats-Unis, au Japon, en Suisse et en Allemagne. Etant donnée l'importance avérée des *O*-fucoses présents sur les récepteurs Notch dans leur activité et l'implication de ces derniers dans la différenciation des cellules myoblastiques, nous avons choisi de nous intéresser à la *O*-fucosylation dans le processus hautement complexe de myogenèse.

Le but de mon projet de thèse était de caractériser les gènes bovins de *O*-fucosyltransférases, de mieux comprendre l'évolution de celles-ci au sein du monde Animal et de déterminer les conséquences d'une modulation de l'expression de ces gènes sur la différenciation de cellules myoblastiques en culture.

Mon travail a débuté par l'identification et la caractérisation fonctionnelle des gènes bovins, *Pofut1* et *Pofut2*, codant les enzymes responsables de la *O*-fucosylation. Leurs organisations génomiques, leurs expressions spatio-temporelles, aussi bien au niveau transcriptionnel que traductionnel ont été déterminées. En parallèle, grâce à la recherche dans les bases de données moléculaires et la sélection de séquences *Pofut* complètes de différents organismes représentatifs du monde vivant, les structures génomiques de *Pofut1* et *Pofut2* d'un nombre conséquent d'espèces ont été déterminées, et une analyse phylogénétique a permis d'étudier leur évolution dans le règne Animal. L'enzyme *Pofut1*, protéine soluble située dans la lumière du réticulum endoplasmique, est une glycoprotéine porteuse de deux *N*-glycanes. Une seconde phase de mon travail a été consacrée à la démonstration de l'importance fondamentale d'un de ces *N*-glycane dans l'activité enzymatique de *Pofut1*. Lorsque j'ai débuté mon doctorat, le gène *Pofut1* codant l'enzyme *O*-fucosyltransférase 1 n'avait été caractérisé que chez l'Homme, la souris et la drosophile, et l'activité de l'enzyme *O*-fucosyltransférase 2, codée par le gène *Pofut2*, n'était pas encore déterminée. Ainsi, non seulement les résultats de mes travaux ont enrichi le répertoire des espèces pour lesquelles les *O*-fucosyltransférases sont désormais connues, permis d'analyser finement la situation chez *Bos taurus* mais ils ont aussi mis en évidence des situations originales, spécifiques à cette espèce. Finalement, j'ai pu commencer l'examen des répercussions d'une surexpression et d'une sous-expression du gène *Pofut1*, sur le déroulement du programme myogénique de cellules myoblastiques bovines en culture primaire et de cellules myoblastiques murines de la lignée C2C12. Pour ce faire, j'ai dû créer et faire usage de nombreux outils nécessaires à l'étude de l'expression et du rôle des *O*-fucosyltransférases dans la différenciation de ces cellules musculaires.

La myogenèse est un processus hautement complexe dont toutes les étapes ne sont pas encore bien comprises. La *O*-fucosylation, de par les protéines qui en sont la cible, comme Notch et ses ligands, doit y jouer un rôle fondamental. Pourtant, cet axe de recherche reste encore aujourd'hui inexploré. Ainsi, je considère que mon étude des *O*-fucosyltransférases ouvre de nouvelles pistes inenvisagées auparavant, riches de découvertes passionnantes qui permettront d'identifier les mécanismes moléculaires gouvernant la différenciation myoblastique. Ce projet s'inscrit donc dans une thématique d'avenir, qui j'en suis convaincue, apportera à la communauté scientifique d'innombrables perspectives de recherche. Le plus beau reste à faire...

Ce manuscrit s'articule autour de trois chapitres.

Le premier chapitre, à caractère bibliographique, s'organise en trois volets. Tout d'abord, nous dressons un état des connaissances sur les deux types de glycosylation des protéines, la N- et la O-glycosylation. Ensuite, une deuxième partie présente les données récentes sur les processus de O-fucosylation sur lesquels porte notre étude. Finalement, après avoir souligné le lien existant entre la glycosylation et la myogenèse, le dernier volet concerne le développement du muscle squelettique.

Le second chapitre, consacré à la présentation des résultats obtenus au cours de ma formation doctorale, s'articule autour de trois parties incluant des publications.

Le dernier chapitre est réservé à la conclusion et aux perspectives envisagées suite à ce travail.

Exposé bibliographique

1 Une modification post-traductionnelle majeure des protéines, la glycosylation

1.1	PRESENTATION ET FONCTIONS DE LA <i>N</i> -GLYCOSYLATION DES PROTEINES	15
1.1.1	<i>Structures et diversité des N-glycanes</i>	15
1.1.2	<i>Biosynthèse des N-glycanes</i>	18
1.1.3	<i>Rôles biologiques des N-glycanes</i>	22
1.2	PRESENTATION ET FONCTIONS DES DIFFERENTS TYPES DE <i>O</i> -GLYCOSYLATION DES PROTEINES.....	27
1.2.1	<i>La O-glycosylation de type GAG</i>	28
1.2.2	<i>La O-glycosylation de type mucine</i>	30
1.2.3	<i>La O-mannosylation</i>	32
1.2.4	<i>La O-N-acétylglucosaminylation</i>	34
1.2.5	<i>La O-glycosylation</i>	34
1.2.6	<i>La O-fucosylation</i>	36
1.2.7	<i>Les différents types de O-glycanes ont des rôles biologiques communs</i>	36

Depuis leur synthèse par la machinerie ribosomale jusqu'à leur localisation finale, les protéines peuvent subir de multiples modifications post-traductionnelles dont la glycosylation qui va nous intéresser tout au long de ce manuscrit. Les glycoprotéines sont constituées d'une chaîne peptidique, sur laquelle sont greffés des sucres sous la forme de monosaccharide voire de chaîne polysaccharidique encore appelée glycane. Plusieurs résidus d'une même chaîne polypeptidique peuvent être substitués par des sucres. Les glycoprotéines sont rencontrées dans différents compartiments cellulaires (réticulum endoplasmique, appareil de Golgi, etc.) et présentes au niveau de la membrane plasmique, dans la matrice extracellulaire et dans le plasma sanguin. Elles sont largement répandues chez les Eucaryotes (Protistes, Levures, Insectes, Plantes, Mammifères, etc.) et occupent des fonctions variées en tant qu'enzymes, protéines de structure, de réserve ou de reconnaissance. La mise en évidence de glycoprotéines chez les Procaryotes (Bactéries et Archées) est assez récente et de nombreuses structures glycaniques sont maintenant décrites chez des organismes divers incluant de nombreux pathogènes tels que *Helicobacter pylori* (Hooper et Gordon 2001) ou *Porphyromonas gingivalis* (Gallagher *et al.*, 2003).

Deux grandes familles de glycoprotéines ont été définies selon la nature de la liaison covalente établie entre le glycane et la partie peptidique : les *N*- et les *O*-glycoprotéines. Les protéines *N*-glycosylées portent un glycane sur l'atome d'azote de la chaîne latérale d'une asparagine comprise dans un motif consensus de type N-X-S/T (X pouvant être n'importe quel résidu autre que la proline). Les protéines *O*-glycosylées portent un monosaccharide ou un glycane fixé par l'intermédiaire de l'atome d'oxygène de la chaîne latérale d'une sérine ou d'une thréonine. Ces deux types de glycosylation peuvent se retrouver simultanément sur une même protéine. Les récepteurs de la famille Notch qui feront l'objet d'une description plus loin dans le manuscrit, portent à la fois des *N*- et des *O*-glycanes (Ishikawa *et al.*, 2005).

Dans cet exposé, nous nous intéresserons uniquement à la glycosylation des protéines chez les Eucaryotes et particulièrement à celle rencontrée chez les Animaux.

1.1 Présentation et fonctions de la *N*-glycosylation des protéines

Les *N*-glycoprotéines représentent la classe la plus abondante de glycoprotéines (Vijay, 1998). La *N*-glycosylation des protéines chez les Eucaryotes est un processus métabolique hautement conservé. En effet, on retrouve des *N*-glycanes aussi bien chez les organismes unicellulaires (Levures, Trypanosomes, etc.) que chez les organismes pluricellulaires (Plantes, Insectes, Mammifères, etc.) (Spiro, 2002). Pour tous les Eucaryotes, cette modification post-traductionnelle est indispensable au bon fonctionnement des cellules (Kukuruzinska et Lennon-Hopkins, 1999), ainsi qu'à la «viabilité» des organismes. Ainsi, un défaut dans la machinerie de *N*-glycosylation peut conduire, chez l'Homme par exemple, à un syndrome de type CDG (Congenital Disorder of Glycosylation) qui peut entraîner des retards mentaux et développementaux, des problèmes gastro-intestinaux ou immunitaires (Freeze et Aebi, 2005).

1.1.1 Structures et diversité des *N*-glycanes

En principe, les glycoprotéines sont potentiellement d'une diversité presque illimitée. L'origine de cette diversité est aussi bien chimique que biologique. La composante protéique des glycoprotéines varie de par sa taille, sa composition, l'agencement des acides aminés et les modifications qui pourraient s'opérer sur certains d'entre eux. Quand une protéine est modifiée par glycosylation, la partie glucidique peut être de séquence et de longueur variable. La diversité des glycanes est aussi liée au fait que les monosaccharides peuvent se combiner entre eux avec des liaisons variables selon l'anomérisation (α ou β).

Les *N*-glycanes sont des oligosaccharides qui possèdent tous une structure pentasaccharidique commune appelée noyau (ou cœur), $\text{Man}_3\text{GlcNAc}_2$ (trois résidus de mannose et de deux résidus de *N*-acétylglucosamine), un des résidus GlcNAc assurant la

liaison à la chaîne polypeptidique. C'est le développement de techniques comme la H-NMR (résonance magnétique nucléaire du proton) qui a permis de mettre en évidence les différentes structures des N-glycanes rencontrés chez les Eucaryotes. Globalement, on peut les classer en trois grandes catégories en fonction des ramifications s'opérant autour du noyau: oligomannosidique, complexe et hybride (Fig.1.) (Geisow, 1991). La proportion relative de ces différents types de glycanes varie en fonction des espèces ainsi qu'en fonction de l'état physiologique ou pathologique de la cellule (stress, cancer, etc.), du tissu ou de l'individu. Il a par exemple été montré que chez le poisson zèbre *Danio rerio*, il n'existe pas de N-glycanes de type oligomannosidique (Guérardel *et al.*, 2006).

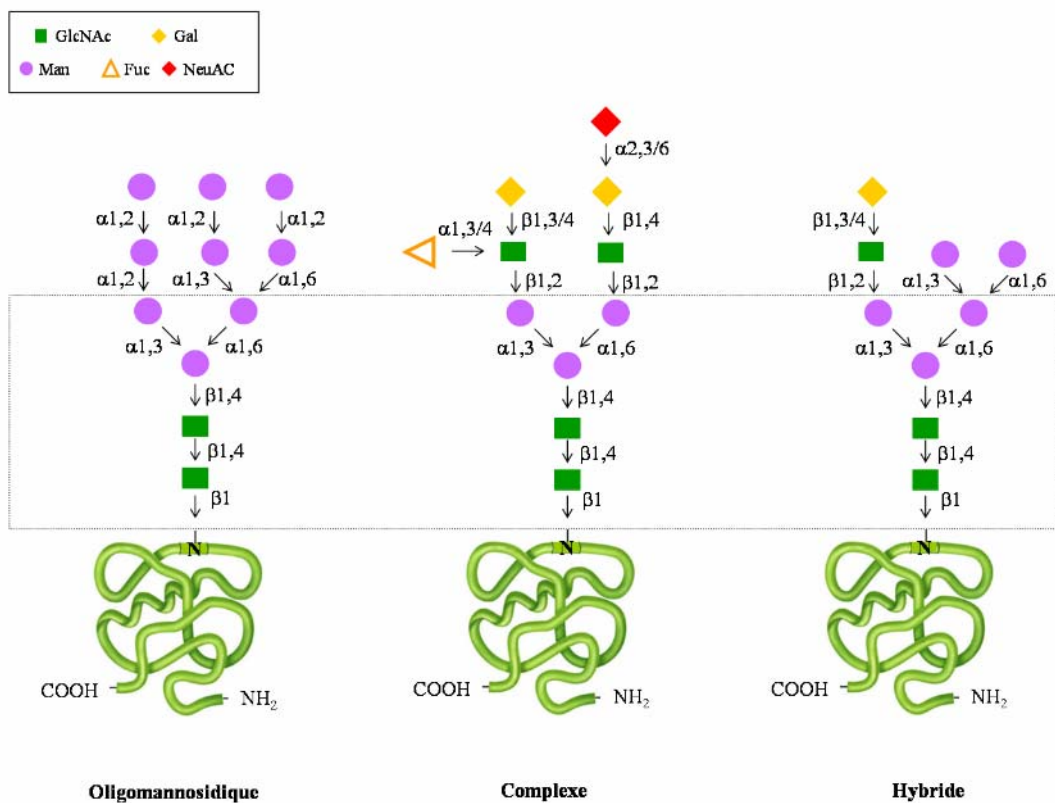


Figure 1 : Exemples de structure des trois classes principales de N-glycanes présentes chez les Eucaryotes.

La partie encadrée correspond au cœur pentasaccharidique commun à tous les N-glycanes (d'après Kornfeld et Kornfeld, 1985).

> Les N-glycanes riches en mannose ou oligomannosidiques sont constitués du noyau, $\text{Man}_3\text{GlcNAc}_2$, additionné de deux à six résidus de mannose. Les levures produisent typiquement ce type de N-glycanes; il y a jusqu'à 100 résidus mannose portés par le noyau

pentasaccharidique (Tanner et Lehle, 1987). Cette hypermannosylation peut engendrer une immunogénicité de la protéine porteuse chez l'Homme. Comparé aux levures, les champignons filamenteux synthétisent le plus souvent des *N*-glycanes avec un nombre relativement faible de résidus mannoses. Aucun oligosaccharide de type complexe contenant de l'acide sialique, du galactose, du fucose et/ou de la *N*-acétylgalactosamine n'a été caractérisé sur les glycoprotéines du groupe des Champignons (Herscovics et Orlean, 1993; Kukuruzinska et Lennon, 1998).

> Les *N*-glycanes complexes possèdent un noyau dont les 2 résidus de mannose sont allongés par une *N*-acétylglucosamine liée en β 1,2. Les antennes des glycanes complexes sont élaborées à partir de ces résidus sur lesquels peut être greffé un résidu de galactose en β 1,3 ou β 1,4. Cette structure GlcNAc β 1,3/4Gal (*N*-acétyllactosamine) peut, dans certains cas, être complexifiée par le greffage d'un fucose sur la GlcNAc en α 1,3 (Mammifères) ou α 1,4 (Mammifères et Plantes) et/ou par le greffage d'un acide sialique en α 2,3 ou α 2,6 sur le galactose (Kornfeld et Kornfeld, 1985). Parfois, le galactose peut-être substitué soit par un fucose en α 1,2 (Mammifères) (Kobata, 1984) soit par un autre galactose en α 1,3 (Mammifères) (Eckhardt et Golstein, 1983). Les *N*-glycanes complexes ne portent pas seulement deux ramifications mais peuvent être tri-, tétra- (Mammifères, Plantes, Insectes), voire penta-antennés (Oiseaux, Poissons) (Sakamoto *et al.*, 2000). De plus, un résidu de fucose peut être lié en α 1,6 (Mammifères et Insectes) ou α 1,3 (Plantes et Insectes) sur la GlcNAc proximale du noyau (Oriol *et al.*, 1999). On peut également trouver un résidu xylose greffé en β 1,2 (Plantes) (Olczak et Watorek, 2000) ou une GlcNAc en β 1,4 (Mammifères et Insectes) (Oguri *et al.*, 1997) sur le premier mannose du noyau. Quelques exemples illustrant ces complexifications sont répertoriés dans la figure 2.

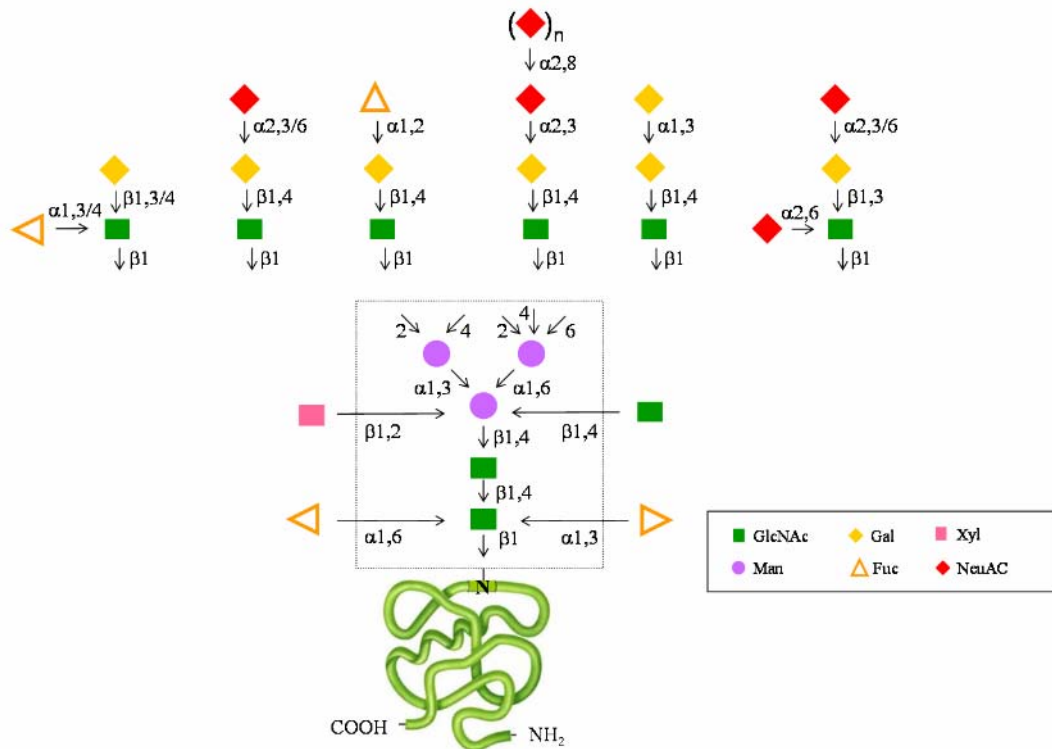


Figure 2 : Différentes ramifications trouvées sur les *N*-glycane complexes.

La partie encadrée correspond au cœur pentasaccharidique commun à tous les *N*-glycane sur lequel cinq ramifications sont possibles.

> Les *N*-glycane hybrides sont des structures intermédiaires entre les glycane oligomannosidiques et complexes. Ils sont constitués de deux antennes saccharidiques de nature différente. L'antenne issue du résidu mannose lié en $\alpha 1,3$ au niveau du noyau est de type complexe alors que celle portée par le mannose lié en $\alpha 1,6$ est de type oligomannosidique (Chen *et al.*, 1998). Le noyau pentasaccharidique peut, quant à lui, subir les mêmes modifications que les *N*-glycane complexes (cf. Fig.2.).

1.1.2 Biosynthèse des *N*-glycane

La translocation co-translationnelle des protéines dans la lumière du réticulum endoplasmique rugueux (RER) constitue la première étape du transport intracellulaire des protéines *N*-glycosylées (Roth, 2002). La présence d'une extension N-terminale appelée peptide signal est nécessaire à la translocation des protéines dans ce système dit de sécrétion (Rapoport *et al.*, 1999). Ce peptide est ensuite clivé par une peptidase après synthèse de la protéine dans la lumière du RER. A leur arrivée dans le réticulum, les protéines subissent de

nombreuses modifications, dont la glycosylation. Une fois leur maturation achevée au sein de l'appareil de Golgi, les glycoprotéines non résidentes de ce compartiment sont triées puis prises en charge par des vésicules golgiennes afin d'assurer leur transport vers leurs zones d'activité (Roth, 2002).

1.1.2.1 Dans le réticulum endoplasmique rugueux

La N-glycosylation débute par la construction d'un oligosaccharide de 14 résidus, de structure $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$, sur un transporteur lipidique activé, le dolicholpyrophosphate (Fig.3.). Les différentes étapes de la biosynthèse de cet oligosaccharide sont catalysées par des glycosyltransférases qui transfèrent successivement chaque ose constitutif (Fig.3. étapes 1, 2 et 4), d'un nucléotide-sucre vers l'accepteur lipidique. La biosynthèse de l'oligosaccharide débute sur la face cytosolique et se poursuit sur la face luminale de la membrane du RER (Abeijon et Hirschberg, 1992), ceci grâce au basculement du dolicholpyrophosphate par un mécanisme de flip-flop (Vishwakarma et Menon, 2005) (Fig.3. étape 3). Le glycane est ensuite transféré "en bloc" sur une asparagine spécifique appartenant au motif N-X-S/T de la chaîne polypeptidique en croissance (Fig.3. étape 5) (Pan et Elbein, 1990). Cette réaction est catalysée par un complexe enzymatique membranaire appelé oligosaccharyltransférase (OST) (Silberstein et Gilmore, 1996). Le dolicholpyrophosphate libéré est alors recyclé sous l'action d'une phosphatase après un nouveau basculement (Fig.3. étapes 6 et 7).

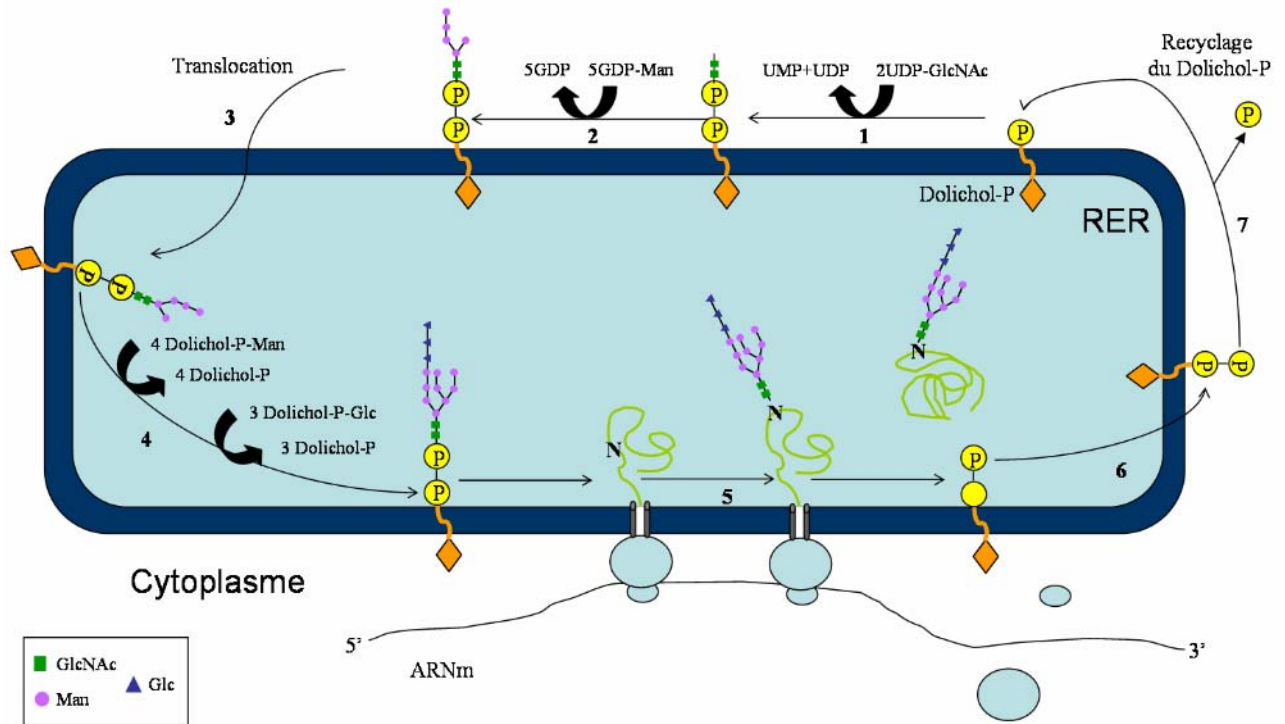


Figure 3 : Voie de biosynthèse du précurseur oligosaccharidique des *N*-glycanes dans le RER. Les enzymes impliquées dans les différentes étapes sont : 1, β 4GlcNAcT; 2, Mannosyl-TI à IX; 4, α 1,3Glc-TI et II; 5, complexe oligosaccharyltransférase; 6, Flippase; 7, Pyrophosphatase.

Les glycanes ainsi *N*-liés à la protéine subissent des modifications séquentielles de leur structure initiale, par l'action de diverses glycosidases et glycosyltransférases, au cours de leur parcours dans le RER puis dans l'appareil de Golgi (Fig.4.). Cette maturation débute par l'élimination du résidu glucose terminal puis des deux résidus glucose restants. Ces réactions sont catalysées respectivement par les glucosidases I et II (Kornfeld et Kornfeld, 1985). La déglucosylation qui aboutit à un glycanes de type oligomannosidique, pourrait débuter avant même que la traduction de la glycoprotéine naissante ne soit achevée.

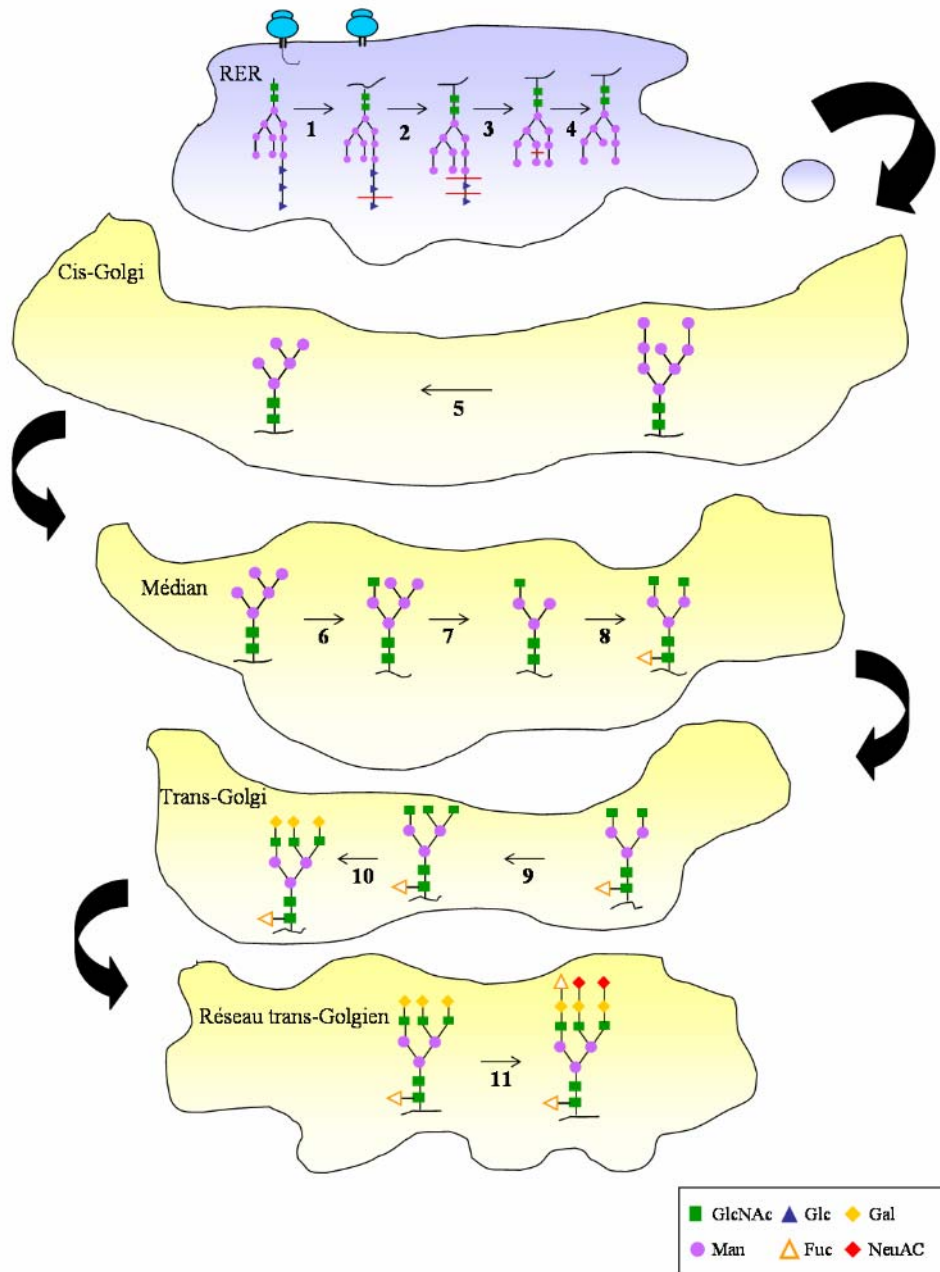


Figure 4 : Maturation des N-glycanes dans le réticulum endoplasmique rugueux et l'appareil de Golgi.

Les enzymes impliquées dans les différentes étapes sont : 1, complexe oligosaccharyltransférase; 2, glucosidase I; 3, glucosidase II; 4, α 2-mannosidase; 5, mannosidase I; 6, β 1,2-N-acétyl-glucosaminyltransférase I; 7, mannosidase II; 8, N-acétyl-glucosaminyltransférase II et fucosyltransférases; 9, N-acétyl-glucosaminyltransférase IV; 10, galactosyltransférases; 11, sialyltransférases et fucosyltransférases.

Il est important de noter que les glycanes des protéines résidentes du RER comme la calréticuline, une protéine chaperonne, subissent peu de modifications et restent de type oligomannosidique (Pagny *et al.*, 2000). A l'inverse, les glycoprotéines qui sont destinées à

rejoindre un autre compartiment cellulaire ou à être sécrétées, transitent par l'appareil de Golgi où leurs glycanes sont modifiés en structures plus complexes et variées.

1.1.2.2 Dans l'appareil de Golgi

Les *N*-glycoprotéines non résidentes du RER ou celles qui doivent y revenir par un mécanisme rétrograde (Mancias et Goldberg, 2005) gagnent alors l'appareil de Golgi par un transport vésiculaire. Pour cela, elles interagissent avec la protéine membranaire ERGIC 53 qui reconnaît les résidus mannoses (Helenius et Aebi, 2001). La maturation des *N*-glycanes se poursuit alors au cours du cheminement de la glycoprotéine dans les différents compartiments de l'appareil de Golgi dans une direction cis-trans (Fig.4.). Leurs structures glycaniques hautement mannosylées sont élaguées sous l'action de la mannosidase I, générant différentes structures de 8 à 5 mannoses. Ces derniers constituent les précurseurs à partir desquels vont être élaborés les *N*-glycanes complexes et hybrides, sous l'action de diverses glycosidases et glycosyltransférases golgiennes.

1.1.3 Rôles biologiques des *N*-glycanes

De multiples fonctions biologiques sont attribuées aux *N*-glycanes selon la localisation de la glycoprotéine à laquelle ils sont associés. A l'intérieur de la cellule, les oligosaccharides favorisent, par exemple, le repliement correct et l'adressage des glycoprotéines qui les portent (Roth, 2002). Les glycoprotéines qui se trouvent à la surface des cellules sont typiquement ce que l'on appelle des protéines de communication. Par le biais de leurs glycanes, elles permettent aux cellules d'interagir avec les différents acteurs de leur environnement : autre cellule, transporteur, anticorps, récepteur, hormone, enzyme, toxine, bactérie et virus (Kukuruzinska et Lennon, 1998). Les rôles des *N*-glycanes étant divers et variés, les exemples cités ci-dessous ne cherchent pas à en faire une liste exhaustive, mais seulement à illustrer leurs rôles principaux.

1.1.3.1 Vis-à-vis de la *N*-glycoprotéine

> Le « contrôle qualité » du repliement des *N*-glycoprotéines

Les glycoprotéines nouvellement synthétisées subissent un contrôle de qualité permettant de vérifier leur conformation. Si celle-ci n'est pas correcte, s'engage alors un

processus de dégradation de la protéine alors qu'elle vient juste d'être synthétisée. Ce mécanisme qui se déroule dans le RE est appelé ERAD (ER-associated degradation pathway) (Hampton, 2002; Kostova et Wolf, 2003). Il est réalisé grâce aux signaux portés à la fois par la chaîne peptidique et le ou les glycane(s) (Fig.5.). Les segments peptidiques hydrophobes de la protéine non repliée sont reconnus par une Hsp70 (Heat shock protein 70 KDa) réticulaire, la chaperonne BiP (Binding Protein) aussi appelée Grp78 (glucose regulated protein-78 KDa). Les glycanes sont, quant à eux, reconnus par des protéines chaperonnes de type lectine. Parmi celles-ci, la calnexine (CNX) et la calréticuline (CRT) se lient à des glycanes monoglucosylés produits transitoirement, par des réactions de glycosylation/déglucosylation, lors du processus de maturation des N-glycanes dans le RER (Hebert *et al.*, 1995). La protéine ERp57 assure, s'il y a lieu, la formation de ponts disulfures (Oliver *et al.*, 1997). Ces mécanismes de reconnaissance sont responsables de la rétention transitoire de la glycoprotéine dans le RER et permettent d'assurer son repliement correct.

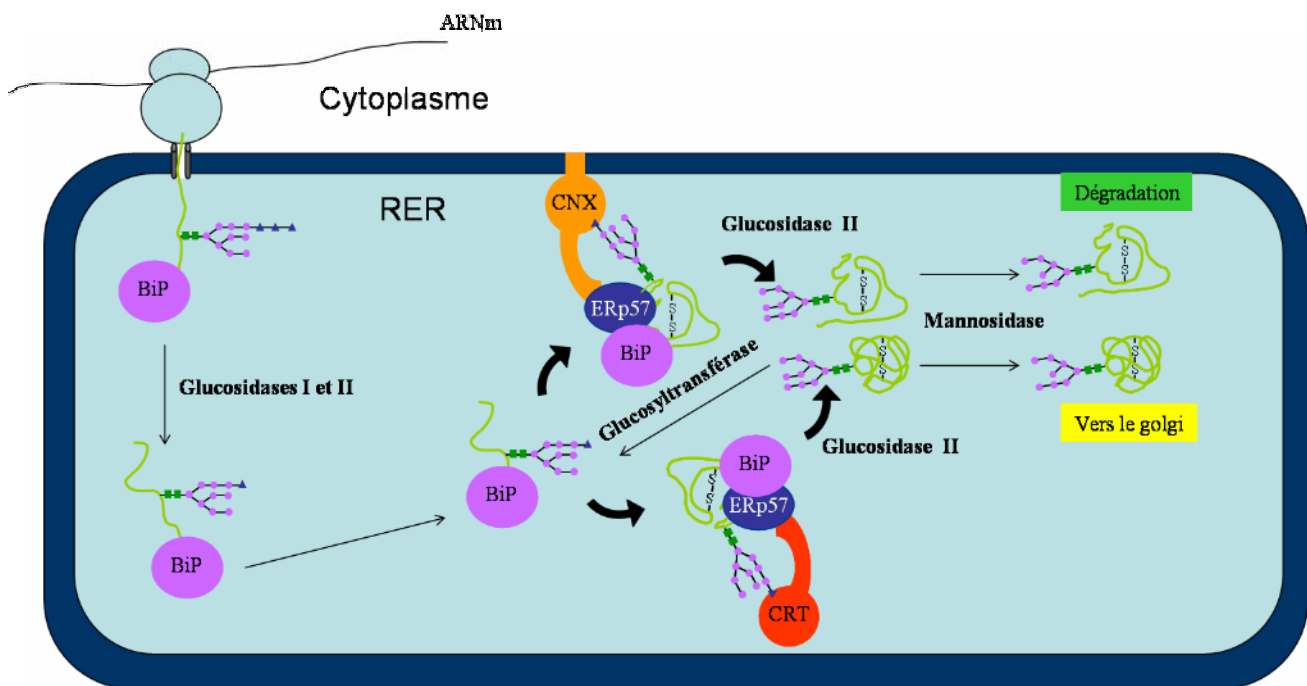


Figure 5 : Repliement d'une protéine glycosylée dans le RER. Cycle de déglucosylation/re-glucosylation.

A son entrée dans le RER, la protéine naissante est prise en charge par BiP puis par CNX et/ou CRT (ERp57 assurant la formation de ponts disulfures) quand deux des résidus glucose ont été enlevés sous l'action des Glucosidases I et II. Quand le dernier résidu glucose est enlevé par la Glucosidase II, les complexes se dissocient. Si la protéine est mal conformée, elle est re-glucosylée, puis le cycle recommence. Au bout d'un certain nombre de cycles, une $\alpha 2$ -mannosidase enlève un résidu mannose; si la protéine est toujours incorrectement repliée, elle est transportée vers le cytoplasme pour sa dégradation. Si sa conformation est correcte, elle quitte le RER pour l'appareil de Golgi (Lehrman, 2001).

Les oligosaccharides riches en mannose retrouvés libres dans le cytosol des cellules attestent du devenir des glycoprotéines nouvellement synthétisées mais mal repliées. En effet, si une protéine est mal conformée, son ou ses glycanes vont être clivés par une *N*-glycanase qui coupe entre l'asparagine de la chaîne polypeptidique et le premier résidu GlcNAc. Ils sont ensuite clivés par d'autres enzymes dans le cytoplasme de la cellule jusqu'à l'hexasaccharide, Man₅GlcNAc₁. Celui-ci sera ultérieurement dégradé en monosaccharides dans les lysosomes par des carbohydrases (Cacan et Verbert, 1999). La protéine sera elle-même dégradée dans les lysosomes ou par le protéasome.

Les glycoprotéines correctement repliées peuvent alors continuer leur chemin vers l'appareil de Golgi. Celles possédant des signaux de rétention du RE et notamment une séquence peptidique C-terminale de type KDEL (ou KKXX, RR) (Monnat *et al.*, 2000) ou une séquence interne RXR (Scott *et al.*, 2003) resteront dans le réticulum ou y reviendront après un bref passage dans le *Cis*-Golgi par un mécanisme de transport rétrograde (Pelham, 1996). Nous reviendrons sur le mécanisme de rétention réticulaire dans le second chapitre de cet exposé.

> Le maintien de la structure des *N*-glycoprotéines et leur oligomérisation

Les glycanes participent au maintien de la structure de la protéine qui les porte. Ils permettent l'adoption d'une conformation biologiquement active, ceci par des liaisons secondaires de type glycane-glycane et/ou glycane-protéine, mettant en jeu des liaisons hydrogènes, ioniques ou des interactions hydrophobes. La protéine s'en trouve ainsi stabilisée et son activité assurée. Par exemple, l'absence de *N*-glycosylation de l' α 1,3-fucosyltransférase IV de rat abolit son activité enzymatique (Baboval *et al.*, 2000). En induisant une conformation spatiale spécifique et de par leur caractère hydrophile, les *N*-glycanes interviennent également dans la solubilité et l'oligomérisation de nombreuses protéines, particulièrement dans les milieux aqueux (Lis et Sharon, 1993).

> La protection des *N*-glycoprotéines vis-à-vis de la protéolyse

Les sites de *N*-glycosylation sont souvent situés dans des régions en feuillets β exposées aux protéases. Les *N*-glycanes, de par l'encombrement stérique qu'ils occasionnent, protègent ces régions de la protéolyse (Kundra et Kornfeld, 1999), notamment durant le transport des glycoprotéines depuis le RER jusqu'à leur destination finale (Wormald et Dwek,

1999). La glycosylation est par conséquent un facteur influençant la durée de demi-vie des glycoprotéines.

> L'adressage de certaines N-glycoprotéines

Les N-glycanes participent également à l'adressage et au devenir des protéines puisqu'ils concourent à leurs localisations intra- ou extra-cellulaires. Par exemple, l'implication directe des N-glycanes a été mise en évidence dans le processus d'adressage des enzymes lysosomales. En effet, il a été démontré que les hydrolases acides destinées aux lysosomes sont des glycoprotéines dont le résidu mannose terminal est phosphorylé en position 6 (étiquette Man-6-P). Cette étiquette constitue un signal permettant le transport de la glycoprotéine de l'appareil de Golgi vers les lysosomes (Dahms et Kornfeld, 1989).

1.1.3.2 Vis-à-vis de l'organisme

De nombreuses études ont mis en évidence le rôle des N-glycanes dans les processus développementaux. A titre d'exemple, les souris KO pour le gène *Mgat1* qui code la GlcNAc transférase I (GlcNAc-TI), enzyme impliquée dans la première étape de la formation des N-glycanes complexes, meurent *in utero* au bout du dixième jour suite à des malformations du tube neural et à un mauvais établissement des axes embryonnaires (Metzler *et al.*, 1994). Des souris KO pour un des gènes codant une β 1,4-N-acétylglucosamine- β -galactosyltransférase (β GlcNAc- β 1-GalT2) assurant le transfert du galactose sur le résidu GlcNAc terminal des N-glycanes complexes et hybrides, présentent un important retard de croissance (Asano *et al.*, 1997).

Un certain nombre de maladies sont associées à des déficiences ou défauts dans le métabolisme des glycoprotéines (Jaeken, 2003; Marquardt et Denecke, 2003). Chez l'Homme, les CDG (Congenital Disorder of Glycosylation) représentent un groupe de maladies innées affectant le processus de N-glycosylation (Freeze et Aebi, 2005). Dès la découverte des premiers déficits enzymatiques, les glycobiochimistes ont proposé une classification tenant compte des grandes étapes de la synthèse des différents types de glycanes. Ainsi, si ces maladies sont la conséquence de mutations au niveau de gènes dont les produits sont impliqués dans la biosynthèse ou le transfert du précurseur oligosaccharidique, on parle de CDG de type I, et si ce sont des gènes dont les produits assurent la maturation des N-glycoprotéines qui sont touchés, on parle de CDG de type II (Dennis *et al.*, 1999). Pour

l'instant, 12 CDG de type I et 6 CDG de type II ont été identifiés (Freeze et Aebi, 2005). Bien que la présentation clinique varie d'un type de CDG à l'autre, une atteinte neurologique sévère associée à un retard mental et à une hypotonie axiale est présente le plus souvent, à laquelle peuvent s'associer des atteintes multiviscérales graves (Jaeken et Matthijs, 2001).

A titre d'exemple, la mucopolysaccharidose II (*I-cell disease*) a servi de « déclencheur » car elle associe un déficit enzymatique à une fonction essentielle de la N-glycosylation, l'adressage. Dans cette maladie, la majorité des hydrolases sont absentes des lysosomes mais sont retrouvées dans le milieu extracellulaire. On sait maintenant que cette affection est due à un déficit en GlcNAc-1-phosphotransférase, normalement présente dans l'appareil de Golgi (Hickman et Neufeld, 1972). Cette enzyme déclenche la synthèse du mannose 6-phosphate (Man-6-P) des chaînes N-glycaniques, reconnu ensuite par son récepteur localisé dans l'appareil de Golgi. Cette lectine adresse alors la N-glycoprotéine portant le Man-6-P vers le lysosome. Dans le cas de la mucopolysaccharidose, l'absence du Man-6-P ne permet plus le transport des hydrolases vers les lysosomes; elles transitent dans l'appareil de Golgi, puis sont sécrétées dans le milieu extracellulaire. Les symptômes de cette maladie sont multiples, hypertrophie gingivale, macroglossie, traits épais, hirsutisme, hernies, peau infiltrée, limitations articulaires, dysostose multiple, hépatomégalie, opacités cornéennes, surdité, retard psychomoteur, nanisme. Le LAD II (*leukocyte adhesion deficiency*) a permis d'illustrer l'importance des motifs saccharidiques dans l'interaction avec les sélectines. Les patients atteints de cette affection souffrent d'infections bactériennes récurrentes associées à une leucocytose périphérique, en relation avec une anomalie d'adhérence des leucocytes circulants (Anderson, et Springer, 1987). Les cellules de ces patients produisent des N-glycoprotéines à faible teneur en fucose dû à une mutation sur le gène codant le transporteur de fucose golgien, ce qui entraîne un déficit en antigènes des groupes sanguins du système ABO et des marqueurs tissulaires de type Lewis, ces derniers étant des ligands de sélectines requis dans la diapédèse des leucocytes.

1.2 Présentation et fonctions des différents types de O-glycosylation des protéines

Contrairement à la N-glycosylation qui débute par l'assemblage d'un oligosaccharide dans le RER qui sera ensuite transféré en bloc sur la protéine naissante, la O-glycosylation se fait par les additions successives de monosaccharides sur un acide aminé hydroxylé d'une protéine. Elle est initiée dans la lumière du RER et se poursuit dans les compartiments de l'appareil de Golgi. Dans une grande majorité des cas, l'acide aminé impliqué est un résidu sérine ou thréonine, mais on rencontre également l'hydroxyproline et l'hydroxylysine (dans le cas du collagène, Yamauchi *et al.*, 1982) ou la tyrosine (dans le cas de la glycogénine, Kao *et al.*, 1993). Nous ne ferons pas ici état de ces deux derniers types de O-glycosylation.

Les principaux monosaccharides constituant les O-glycanes sont la N-acétylgalactosamine, la N-acétylglucosamine, le galactose, le glucose, le xylose, le mannose, le fucose et l'acide sialique. Alors que pour les N-glycoprotéines, la liaison entre glycanes et squelette peptidique est unique, puisqu'elle lie une asparagine à une GlcNAc, il existe différents types de liaisons pour les O-glycoprotéines en fonction de la nature du glycanes et de l'acide aminé engagé (Fig.6.).

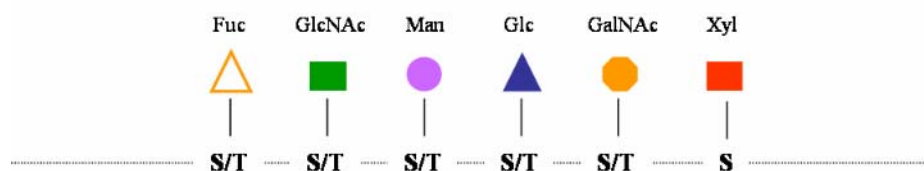


Figure 6 : Représentation schématique des principaux monosaccharides O-liés sur des résidus Sérine et Thréonine.

Les signaux qui déterminent quel acide aminé sera O-glycosylé ne sont pas bien connus. En effet, le site de O-glycosylation ne fait pas forcément intervenir un acide aminé

compris dans une séquence peptidique consensus, sauf dans les cas de la *O*-glucosylation et la *O*-fucosylation.

La formation du *O*-glycane, après ajout du premier résidu, se poursuit généralement par les additions successives d'autres monosaccharides (Brockhausen *et al.*, 1995). Les chaînes glycaniques ainsi formées peuvent être ramifiées et relativement courtes (moins de 15 résidus dans le cas des *O*-glycanes de type mucine) ou non ramifiées et de longueurs variables (à titre d'exemple, de 1 à 4 résidus pour la *O*-fucosylation ou la *O*-mannosylation et jusqu'à 25000 résidus dans le cas des glycosaminoglycanes des protéoglycanes).

A la différence des *N*-glycanes qui présentent dans leurs structures et leurs rôles biologiques de nombreux points communs, il est difficile de définir des caractéristiques communes aux *O*-glycanes. En effet, les structures *O*-glycaniques sont extrêmement variées et présentent de nombreuses différences, aussi bien au niveau de leurs synthèses qu'au niveau de leurs rôles biologiques. Dans le domaine de la Glycobiologie, la *O*-glycosylation est un thème de recherche relativement récent (fin des années 70) par rapport à l'intérêt porté à la *N*-glycosylation (début des années 60). Les données disponibles sur les *O*-glycanes sont donc beaucoup moins conséquentes. De nombreuses *O*-glycosyltransférases restent encore à caractériser aussi bien d'un point de vue génétique que biochimique, et les rôles de nombreuses structures *O*-glycaniques doivent encore être déterminés.

Ainsi, cette partie concernant la *O*-glycosylation est organisée de façon à souligner cette diversité. En effet, chaque sous-partie est dédiée à la description d'un type de *O*-glycosylation donné et à la présentation de son rôle biologique au travers d'exemples pertinents.

1.2.1 La *O*-glycosylation de type GAG

Les glycosaminoglycanes (GAG) sont de longues chaînes non ramifiées d'unités disaccharidiques répétées, généralement un sucre aminé (*N*-acétylglucosamine ou *N*-acétylgalactosamine) et un acide uronique (glucuronique ou iduronique). Elles sont liées de façon covalente à une protéine par l'intermédiaire d'un résidu sérine et forment ainsi les protéoglycanes. Les types de répétition définissent quatre familles principales de protéoglycanes : les chondroïtines sulfates, les dermatanes sulfates, les héparanes sulfates et les kératanes sulfates. Mises à part quelques chaînes de kératane sulfate qui sont liées à la

protéine par une liaison *N*-glycosidique, la liaison des GAG aux protéines fait intervenir le tétrasaccharide : GlcA β 1,3Gal β 1,3Gal β 1,4Xyl β 1-O-S (Fig.7.) (Fondeur-Gelinotte *et al.*, 2006) sur lequel sont ensuite fixés les autres sucres dans l'appareil de Golgi. Les protéoglycanes sont des constituants majeurs de la membrane plasmique, du glycocalyx et de la matrice extracellulaire (MEC) (Vidal y Plana, et Karzel, 1980). La partie glucidique peut représenter jusqu'à 95% de la masse moléculaire d'un protéoglycane, alors qu'elle ne constitue que 1 à 60% de la masse moléculaire d'une glycoprotéine.

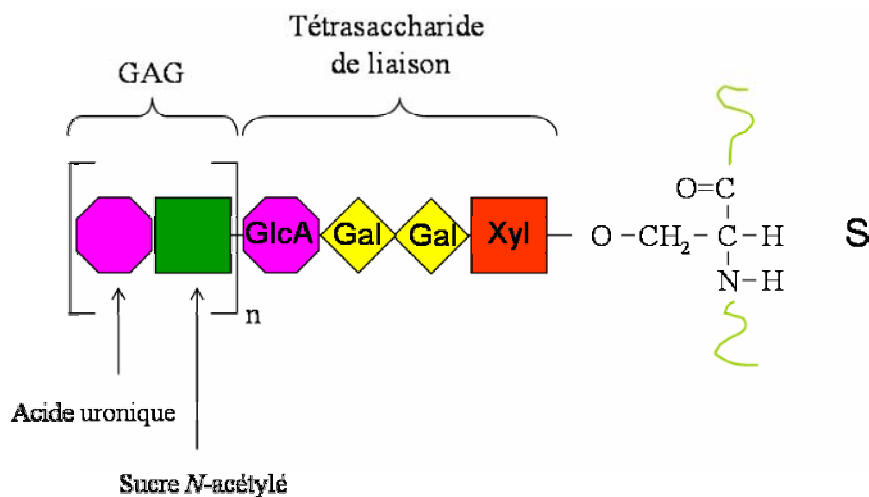


Figure 7 : Représentation schématique d'un protéoglycane.

La taille de la protéine cœur (20 à 200 KDa), le nombre, la longueur et la nature des chaînes glycaniques ajoutées aux modifications que peuvent subir les GAG (*N*-sulfatation, *N*-déacétylation) sont une source de complexité et de combinaison quasi infinie.

Dans la MEC, les protéoglycanes forment un gel hydraté dont la taille des pores et la densité des charges varient selon la composition moléculaire des GAGs. Ils jouent ainsi un rôle dans la viscosité et la résistance aux compressions des tissus (Schonherr *et al.*, 1995). Ils facilitent la migration des cellules au cours de la morphogenèse (Sanes, 2003). Lors des processus tumoraux, ils sont impliqués dans la modulation des interactions cellulaires (Ma et Geng, 2000) et dans la formation des plaques amyloïdes dans les pathologies comme la maladie d'Alzheimer (Gupta-Bansal *et al.*, 1995). L'existence d'un tel gel contrôle la vitesse de diffusion ainsi que la disponibilité des molécules hydrosolubles (Kolset et Gallagher, 1990). Les protéoglycanes de la MEC peuvent donc interagir avec de nombreux facteurs de croissance (fibroblast growth factor, wingless, transforming growth factor, etc.) qui initient

les cascades de signalisation intracellulaire (Haltiwanger et Lowe, 2004). Enfin, leurs multiples fonctions peuvent être exploitées à des fins thérapeutiques. Par exemple, l'héparine est depuis longtemps utilisée pour son activité anticoagulante (Petitou *et al.*, 1988). De même, les chondroïtines sulfates sont utilisées pour ralentir voire stopper le développement de l'arthrose (Leeb *et al.*, 1996).

1.2.2 La O-glycosylation de type mucine

La O-glycosylation de type mucine est initiée par une famille d'enzymes localisées dans l'appareil de Golgi, les *N*-acétylgalactosamine:polypeptide *N*-acétylgalactosaminyltransférases (ppGaNTases). Elles catalysent le transfert d'un résidu GalNAc à partir d'un donneur activé, l'UDP-GalNAc, sur les résidus S ou T de protéines transmembranaires ou sécrétées. Chez les Mammifères, 12 membres appartenant à cette famille de glycosyltransférases ont pour l'heure été identifiés; ce nombre étant sans doute sous-estimé puisque des recherches *in silico* montrent qu'il en existerait plus de 24 chez l'Homme (Ten Hagen *et al.*, 2003). Le résidu GalNAc peut rester sous la forme de monosaccharide ou être allongé par d'autres enzymes (Fig.8.). Un glycanes de type mucine peut en effet contenir jusqu'à 15 sucres.

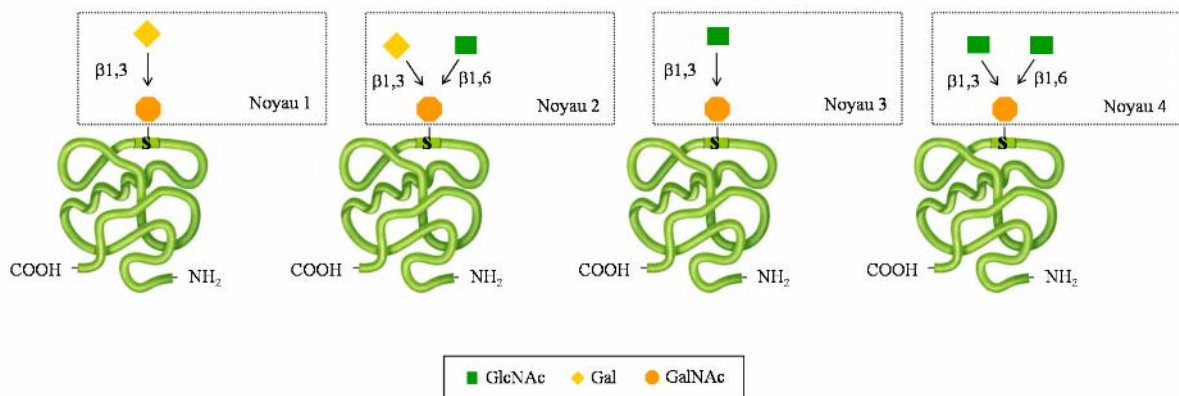


Figure 8 : Représentation des quatre principaux noyaux des O-glycanes de type mucine.

Le résidu GalNAc peut-être allongé par un Galactose en β1,3 (noyau 1) et une GlcNAc en β1,6 (noyau 2), ou une GlcNAc en β1,3 (noyau 3) et une GlcNAc en β1,6 (noyau 4). Ces ensembles peuvent également être complexifiés par l'addition de GalNAc, GlcNAc, NeuAc et Fuc (Haltiwanger et Lowe, 2004).

Les O-glycanes de type mucine se retrouvent majoritairement sur des protéines de l'épithélium, de l'endothélium vasculaire et à la surface des globules rouges. Ces structures définissent différents types d'antigène. Elles peuvent constituer les déterminants des groupes sanguins (ABO) et tissulaires (Lewis) (Van den Steen *et al.*, 1998). Dans ces deux cas, les structures glycaniques sont également portées par des lipides, mais nous n'en ferons pas état ici. Les antigènes Tn présents à la surface de nombreuses cellules cancéreuses sont aussi des structures glycaniques de type mucine (Kannagi *et al.*, 1989).

Cette modification glycosidique doit son nom à sa présence sur les mucines qui sont des composants glycoprotéiques majeurs des muqueuses. Leur O-glycosylation contribue à leur importante viscosité et à leur capacité à intervenir dans les interactions cellule-cellule (Carraway et Hull, 1991; Strous et Dekker, 1992). Les mucines ont un rôle protecteur contre les infections (l'expression des gènes de mucines serait induite lors de l'infection par des micro-organismes (Dohrman *et al.*, 1998)), la déshydratation, les agressions physiques et chimiques des cellules, tout en permettant une diffusion facilitée des molécules hydrosolubles (Perez-Vilar et Hill, 1999). Outre leurs rôles protecteurs, chez certains parasites comme *Trypanosoma cruzi*, agent de la maladie de Chagas chez l'Homme, les mucines seraient nécessaires à l'invasion des cellules hôtes (Pollevick *et al.*, 2000).

Dans le but d'appréhender le rôle de la O-glycosylation de type mucine, des invalidations de gènes codant des ppGaNTases ont été effectués chez la souris. La suppression des gènes *ppGaNTase-TIV* et *V* (résultats non publiés cités dans Ten Hagen *et al.*, 2003) ou du gène *ppGaNTase-TXIII* (Hennet *et al.*, 1995) a mis en évidence des dérèglements de la O-glycosylation dans certains types cellulaires (cellules de Purkinje, lymphocyte B). Néanmoins, les souris KO ne présentent aucun phénotype particulier et sont donc indiscernables des souris sauvages (Zhang *et al.*, 2003). Ceci serait sans doute dû à la redondance des activités des nombreuses ppGaNTases. Cependant, des études effectuées chez la drosophile montrent que les O-glycanes de type mucine seraient impliqués dans le développement embryonnaire. Peu de données sont disponibles à l'heure actuelle sur ces structures glycaniques chez cet organisme. *Drosophila melanogaster* possède 15 gènes codant des ppGaNTases et la mutation de l'un d'entre eux (*pgant35A* ou *l(2)35A*) aboutit à l'arrêt du développement post-embryonnaire au stade pupal (Schwientek *et al.*, 2002).

1.2.3 La O-mannosylation

Avant la mise en évidence de cette modification chez les Mammifères, il y a moins d'une trentaine d'années (Finne *et al.*, 1979), on la croyait seulement présente chez la levure et quelques autres champignons où elle intervient dans la sécrétion des protéines, le maintien de l'intégrité cellulaire et la sporulation (Strahl-Bolsinger *et al.*, 1999). Depuis, de nombreuses structures ont été trouvées chez les Mammifères, la plupart d'entre-elles comportant le «core glycan», Gal β 1,4GlcNAc β 1,2Man1-O-S/T (Haltiwanger et Lowe, 2004).

La synthèse d'un O-mannosylglycane débute dans le RER par le greffage, sur un résidu S ou T d'une protéine en élongation, d'un mannose par l'intermédiaire du dolicholphosphate activé. L'allongement se poursuit par les additions successives d'autres monosaccharides dans l'appareil de Golgi.

Cette modification est retrouvée sur des glycoprotéines du cerveau, du système nerveux et des muscles squelettiques. Environ 30% des protéines O-glycosylées présentes dans le cerveau des Mammifères seraient O-mannosylées (Chai *et al.*, 1999). Des O-glycane de structure NeuAc α 2,3Gal β 1,4GlcNAc β 1,2Man1-O-S/T (Fig.9.) ont été mis en évidence sur des domaines peptidiques riches en résidus S et T de l' α -dystroglycane musculaire bovin (Chiba *et al.*, 1997). L' α -dystroglycane est une glycoprotéine extracellulaire de 156 KDa. Par l'intermédiaire de ses O-mannosylglycane, elle se fixe à la laminine, une protéine de la lame basale assurant ainsi la liaison entre l'actine du cytosquelette et la MEC. Comme nous le verrons plus loin dans cet exposé, son hypomannosylation chez l'Homme et la souris, entraîne la survenance de dystrophies musculaires.

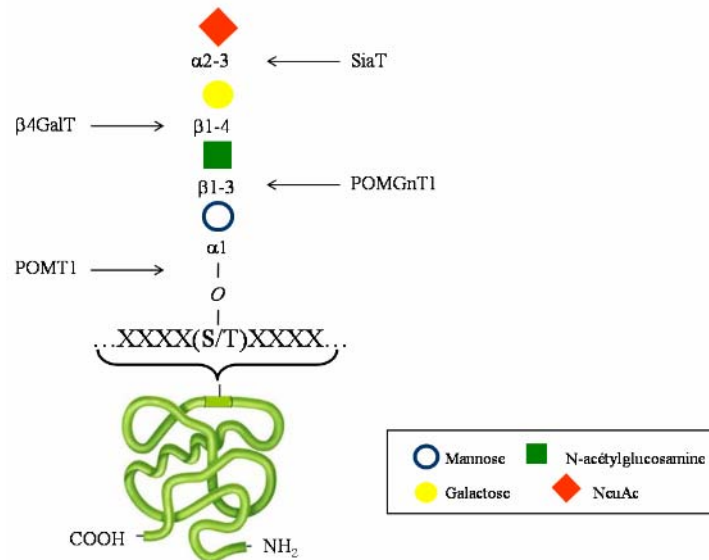


Figure 9 : Représentation schématique du O-mannosylglycane présent sur l' α -dystroglycane. Le greffage du O-mannose est effectué par l'enzyme POMT1 ou Protéine O-mannosyltransférase 1. Ce monosaccharide est ensuite allongé en tétrasaccharide par les actions séquentielles de POMGnT1 (Protéine O-mannoside-N-acétylglucosaminyltransférase), une $\beta 1,4$ -galactosyltransférase et une $\alpha 2,3$ -sialyltransférase. Pour l'instant, aucun motif consensus pour cette modification post-traductionnelle n'a été déterminé.

Les gènes codant les enzymes responsables des premières étapes de la formation du O-mannosylglycane ont été identifiés. Le gène *POMT1*, codant une O-mannosyltransférase, est l'homologue du gène *rt* (*rotated abdomen*) de drosophile (Jurado *et al.*, 1999). Des embryons mutants pour ce gène sont viables mais souffrent d'importants problèmes développementaux : défauts dans la formation des muscles, mauvaise acquisition de la polarité ventro-latérale, rotation du corps au niveau de l'abdomen (Martin-Blanco et Garcia-Bellido, 1996). Chez l'Homme, *POMT1* est exprimé dans de nombreux tissus et code une protéine transmembranaire du RE possédant de nombreuses homologies avec les mannosyltransférases de levures. Le gène *POMGnT1* code la O-mannoside-N-acétylglucosaminyltransférase 1 responsable du greffage d'un résidu GlcNAc sur le monosaccharide $\text{Man}\alpha 1\text{-S/T}$ (Yoshida *et al.*, 2001; Zhang, Betel *et al.*, 2002). L'allongement en tri- et tétrasaccharide se fait grâce à l'ajout d'un résidu galactose puis d'un acide sialique. Les enzymes responsables de ces deux derniers greffages ne sont pas encore identifiées.

1.2.4 La O-N-acétylglucosaminylation

Ce type de O-glycosylation, largement répandu dans les cellules eucaryotes, est catalysé par l'enzyme OGT (*O*-GlcNAc Transférase), une glycosyltransférase présente à la fois dans le noyau et le cytoplasme des cellules (Akimoto *et al.*, 1999). Cette modification par un seul sucre est fréquemment retrouvée sur des protéines appartenant à ces deux compartiments cellulaires (Torres et Hart, 1984) : ARN polymérase II, protéines des pores nucléaires, facteurs de transcription (SP1- specificity protein-1, SRF –Serum response Factor, récepteur aux œstrogènes, etc.), protéines de proto-oncogènes (c-myc, c-fos, c-jun, etc.), protéines de suppresseurs de tumeur (p53), protéines du cytosquelette (Tau, Vinculine, Ankyrine, Clathrine, etc.) (Shafi *et al.*, 2000).

La O-N-acétylglucosaminylation (ou O-GlcNAcylation) est impliquée, par exemple, dans le contrôle de la transcription, l'assemblage des neurofilaments ou la survenue de la maladie d'Alzheimer (Yao et Coleman, 1998). Elle aurait un rôle de régulation similaire à celui de la phosphorylation. Certaines protéines comme l'ARN polymérase II et c-myc peuvent être soit phosphorylées soit O-GlcNAcylées sur le même acide aminé (Hart, 1997). La O-GlcNAcylation, comme la phosphorylation, est réversible. Une N-acétyl- β -D-glucosaminidase a d'ailleurs été mise en évidence dans le noyau et le cytoplasme des cellules eucaryotes (Dong et Hart 1994).

L'enzyme OGT est une protéine très conservée d'un point de vue évolutif puisque 65% de ses acides aminés sont identiques entre l'enzyme humaine et celle de *Caenorhabditis elegans* (Lubas *et al.*, 1997). Un tel degré de conservation souligne une fonction essentielle pour cette enzyme. En effet, une délétion du gène correspondant, provoque l'apoptose des cellules ES (Embryonic Stem) murines (Shafi *et al.*, 2000).

1.2.5 La O-glycosylation

Ce type de modification post-traductionnelle a lieu sur les domaines EGF (Epidermal growth factor) et EGF-like que comportent de nombreuses protéines de surface (Harris et Spellman, 1993). Ce sont des motifs peptidiques répétés de 30 à 40 acides aminés, souvent impliqués dans les interactions entre protéines. Ils doivent leur nom à leur homologie avec la protéine EGF, un facteur de croissance régulant entre autres le taux de calcium intracellulaire.

Ces domaines sont caractérisés par la présence de six résidus cystéines formant trois ponts disulfures impliqués dans le maintien de leur conformation tridimensionnelle. Bien que les boucles peptidiques créées par les ponts disulfures soient plus petites dans les domaines EGF-like que dans les domaines EGF, ils n'en possèdent pas moins une conformation similaire. Dans la suite du manuscrit et pour ne pas alourdir inutilement le texte, nous emploierons uniquement le terme EGF pour désigner ces deux types de domaines.

Le O-glucosylglycane complet (Xyl α 1,3Xyl α 1,3Glc β 1-S/T) et des formes intermédiaires (Xyl α 1,3Glc β 1-S/T et Glc β 1-S/T) ont été trouvés sur des domaines EGF de nombreuses protéines. Initialement identifiés sur les facteurs de coagulation VII et IX bovin (Hase *et al.*, 1988), ces O-glycanes ont ensuite été découverts sur la protéine Z humaine impliquée dans la fécondation (Nishimura *et al.*, 1989), le récepteur Notch et ses ligands (Moloney, Shair *et al.*, 2000) impliqués dans de nombreux processus développementaux, et la thrombospondine, notamment impliquée dans l'adhésion et l'agrégation plaquettaire (Nishimura, Yamashita *et al.*, 1992).

La comparaison des sites de liaison du O-glucose sur ces diverses protéines a permis d'identifier un motif consensus de O-glycosylation. Le glucose est directement fixé sur le groupement hydroxyle d'un résidu sérine ou thréonine appartenant au motif, C¹XS/TXPC², où C¹ et C² sont la première et la deuxième cystéines conservées d'un domaine EGF.

Une forte activité O-glucosyltransférase a été mise en évidence dans des extraits protéiques de cellules CHO (Shao *et al.*, 2002), ceci grâce à des tests d'activité enzymatique *in vitro*, utilisant de l'UDP-glucose radiomarqué comme substrat donneur et le premier domaine EGF du facteur de coagulation VII humain comme substrat accepteur. Une telle activité est également présente dans diverses lignées cellulaires comme SF9 (cellules d'insectes), HD11 (macrophages de poulet), NIH 3T3 (fibroblastes d'embryon murin) et HELA (Shao *et al.*, 2002). Aucune activité O-glucosyltransférase n'a été trouvée chez la levure. Pour l'instant, chez les Animaux, la ou les enzyme(s) responsable(s) de ce type de greffage ne sont pas encore identifiées, contrairement aux Végétaux où plusieurs O-glucosyltransférases ont été caractérisées (Veach *et al.*, 2003).

1.2.6 La O-fucosylation

La O-fucosylation affecte aussi bien les domaines EGF que les domaines TSR (Thrombospondin type I repeat). Ces derniers sont des motifs peptidiques répétés d'environ 60 acides aminés riches en résidus W, S, R et C conservés. Tout comme les domaines EGF, ils sont caractérisés par la présence de six résidus cystéines impliqués dans la formation de trois ponts disulfures. La O-fucosylation fera l'objet d'une présentation détaillée dans le deuxième chapitre de ce manuscrit.

1.2.7 Les différents types de O-glycanes ont des rôles biologiques communs.

Tout comme dans le cas des N-glycanes, les fonctions biologiques des O-glycanes apparaissent comme essentielles aux organismes pluricellulaires, et il serait présomptueux ici de penser pouvoir en faire une synthèse complète.

1.2.7.1 Vis-à-vis de la O-glycoprotéine

> Le maintien de la structure des O-glycoprotéines

La O-glycosylation participe au maintien de la structure tertiaire des O-glycoprotéines. Ainsi, la O-glycosylation du récepteur de la neurotrophine présent à la surface de certaines cellules serait nécessaire à son orientation et à son éloignement de la membrane plasmique, facilitant ainsi l'interaction avec son ligand (Chapman *et al.*, 1996). Les O-glycanes modulent également l'agrégation des protéines en influençant leur structure quaternaire (Van den Steen *et al.*, 1998). Le maintien de l'intégrité structurale est fondamental pour l'activité des enzymes. En effet, la déglycosylation enzymatique par une O-glycanase ou chimique par l'acide trifluorométhanesulfonique de la protéine LPH humaine (Lactase-phlorizine hydrolase), responsable de l'hydrolyse du lactose, diminue sa vitesse de catalyse (Naim et Lentze, 1992).

> La stabilisation thermique des O-glycoprotéines et leur protection vis-à-vis de la protéolyse

Comme la N-glycosylation, la O-glycosylation participe à la stabilisation thermique des glycoprotéines et à leur protection vis-à-vis de la protéolyse. Par exemple, le O-glycane

de la protéine G-CSF (Granulocyte colony stimulating factor) empêche sa dénaturation à 37°C (Oh'eda *et al.*, 1990). De même, une hypo-O-N-acétylglucosaminylation (O-GlcNAc) du facteur de transcription Sp1 entraîne sa rapide dégradation par des protéases (Han et Kudlow, 1997).

> La modulation des interactions hôte/pathogène

Les O-glycanes présents à la surface des cellules sont impliqués dans la modulation des interactions hôte/pathogène. En effet, un certain nombre de microorganismes pathogènes (virus, bactéries et parasites) se lient aux O-glycanes de la cellule hôte et/ou les utilisent afin de ne pas être détectés par le système immunitaire. Ainsi, le virus de l'Influenza se fixe sur une cellule épithéliale respiratoire, par l'intermédiaire d'une de ses glycoprotéines de surface, l'hémagglutinine, qui interagit avec l'acide sialique présent en position terminale des O-glycoprotéines de la membrane de la cellule hôte (Meanwell et Krystal, 2000). Après réplication virale au sein de la cellule et afin de propager l'infection, ces pathogènes sont capables de modifier, par clivage enzymatique, la composition en acide sialique des O-glycanes de la membrane plasmique, ceci par l'intermédiaire d'un autre de leurs antigènes de surface ayant une activité glycosydase, la neuraminidase (Zhang et Xu, 2006).

1.2.7.2 Vis-à-vis de l'organisme

Les O-glycanes constituent des marqueurs du dysfonctionnement cellulaire. En effet, les fractions O-glycaniques des glycoconjugués membranaires de type mucine sont profondément modifiées dans les cellules cancéreuses, entraînant des variations de leur masse moléculaire et de leur taux de sialylation. Cela se traduit souvent par l'apparition ou la prédominance de certains épitopes O-glycaniques comme les antigènes tissulaires de type Le^x (Galβ1,4(Fuca1,3)GlcNAcβ1,3GalNAcαS/T), sLe^x(NeuAcα2,3Galβ1,4(Fuca1,3)GlcNAcβ1,3GalNAcαS/T), Le^y (Fuca1,2Galβ1,4(Fuca1,3)GlcNAcβ1,3GalNAcαS/T) et sLe^a (NeuAcα2,3Galβ1,3(Fuca1,4)GlcNAcβ1,3GalNAcαS/T), dans les cas de cancers du poumon (Jordon *et al.*, 1989), du colon (Cooper *et al.*, 1991) et de l'estomac (Sakamoto *et al.*, 1989). Dans le cas de cancers du colon, les cellules présentent également un taux anormalement élevé de structures de type antigénique Tn (GalNAcαS/T), TF (Galβ1,3GalNAcαS/T) et sialyl-Tn (NeuAcα1,6GalNAcαS/T), alors que la structure, GlcNAcβ1,3GalNAc, normalement exprimée, est absente des cellules cancéreuses (Byrd et Bresalier, 2004).

Le taux anormalement élevé de structures de type Lewis à la surface des cellules cancéreuses serait propice à la formation des métastases. Des études *in vitro* ont démontré que l'attachement des cellules tumorales à l'endothélium, pour le traverser, a lieu par l'intermédiaire d'interactions entre des structures antigéniques sLe^x ou sLe^a et les sélectines E (Tozeren *et al.*, 1995). De même, l'augmentation de la quantité d'antigènes sialyl-Tn dans des cellules de cancer du sein augmenterait leur pouvoir métastatique (Julien *et al.*, 2005). L'importance des antigènes sLe^a et sLe^x dans la régulation de la cascade métastatique a été confirmée par des études cliniques. Une étude réalisée sur un cancer du foie démontre que l'expression de chaque antigène est plus importante dans les lésions métastatiques que dans les tumeurs primaires (Hasegawa *et al.*, 1993).

De par leur localisation, souvent membranaire, et leur importance dans les processus de reconnaissance et de communication cellule-cellule, les O-glycanes sont impliqués dans de nombreux processus développementaux, d'où l'intérêt de l'étude de telles structures. Le second chapitre de cet exposé est consacré aux O-fucosyltransférases, enzymes responsables d'une modification post-traductionnelle qui se produit précocement au cours du développement embryonnaire des Vertébrés.

2 Nos glycosyltransférases d'intérêt, les *O*-fucosyltransférases

2.1	UNE GLYCOSYLTRANSFERASE IMPLIQUEE DANS LE DEVELOPPEMENT EMBRYONNAIRE, LA <i>O</i> -FUCOSYLTRANSFERASE 1	45
2.1.1	<i>Cibles de l'enzyme Pofut1, les EGF</i>	45
2.1.2	<i>Caractérisation de l'enzyme Pofut1</i>	48
2.1.3	<i>Rôles du O-fucose des domaines EGF</i>	55
2.2	DECOUVERTE RECENTE D'UNE ENZYME HOMOLOGUE A POFUT1, LA <i>O</i> -FUCOSYLTRANSFERASE 2.....	63
2.2.1	<i>Une seconde cible de la O-fucosylation, les TSR</i>	63
2.2.2	<i>Caractérisation de l'enzyme Pofut2</i>	65
2.2.3	<i>Rôles du O-fucose des domaines TSR</i>	68
2.3	CARACTERISTIQUES MOLECULAIRES ET FONCTIONNELLES DES <i>O</i> -FUCOSYLTRANSFERASES.....	71
2.4	EXISTE T'IL UNE TROISIEME <i>O</i> -FUCOSYLTRANSFERASE ?	75
2.5	<i>O</i> -FUCOSYLATION ET DEVELOPPEMENT EMBRYONNAIRE.....	77

En 1975, pour la première fois, un O-fucosylglycane de structure Glc β 1,3Fuc α 1-S a été mis en évidence par chromatographie et spectrométrie de masse à partir d'urine humaine (Hallgren *et al.*, 1975). La O-fucosylation est l'ajout d'un fucose sur un résidu sérine ou thréonine (S/T) compris dans deux grands types de domaines : les EGF (Harris et Spellman, 1993) au sein du motif consensus C²X₄₋₅(S/T)C³, où C² et C³ représentent la deuxième et la troisième cystéines du domaine (Shao et Haltiwanger, 2003), et les TSR (Hofsteenge *et al.*, 2001) au sein du motif consensus WX₅C¹X₂₋₃(S/T)C²X₂G (Luo et Haltiwanger, 2005), où C¹ et C² sont la première et la deuxième cystéines du domaine TSR (Fig.10.).

Les domaines EGF et TSR ont la particularité de posséder six résidus cystéines impliqués dans la formation de trois ponts disulfures, leur conférant leur conformation tridimensionnelle.

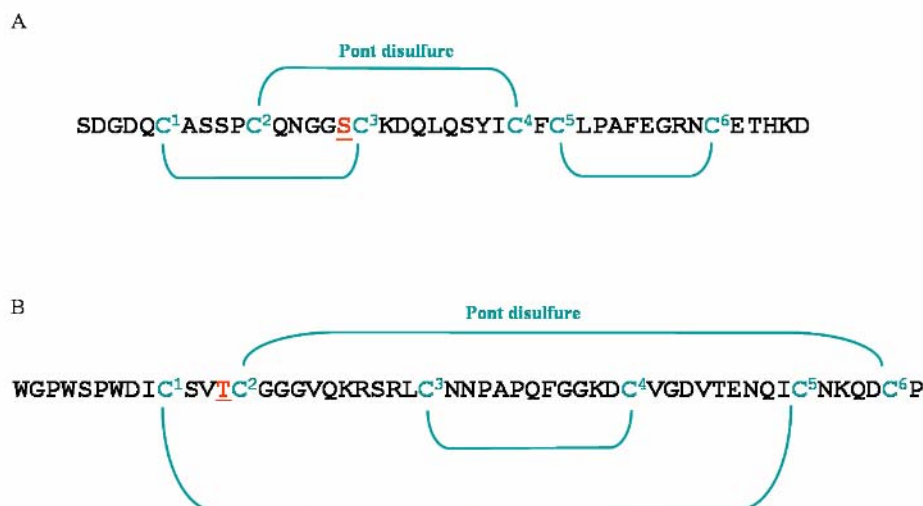


Figure 10 : Représentation des domaines EGF et TSR.

(A) Séquence du domaine EGF du facteur de coagulation IX humain (B) Séquence du domaine TSR3 de la thrombospondine I humaine. Le site de O-fucosylation (S ou T) est noté en rouge, les cystéines conservées et les ponts disulfures en vert.

La O-fucosylation des domaines EGF et TSR dépend respectivement de deux enzymes. Celle responsable de la O-fucosylation des EGF est Pofut1 (Protein O-fucosyltransférase 1) (Wang *et al.*, 1996) alors que celle greffant le fucose sur les TSR est Pofut2 (Protein O-fucosyltransférase 2) (Luo, Koles *et al.*, 2006; Luo, Nita-Lazar *et al.*, 2006).

2.1 Une glycosyltransférase impliquée dans le développement embryonnaire, la O-fucosyltransférase 1

2.1.1 Cibles de l'enzyme Pofut1, les EGF

2.1.1.1 Protéines possédant des EGF O-fucosylés

L'enzyme O-fucosyltransférase 1 ou Pofut1 est responsable du greffage du O-fucose sur certains domaines EGF des protéines. Les premiers domaines EGF démontrés comme O-fucosylés ont été ceux de la Pro-Urokinase (Kentzer *et al.*, 1990) et de l'activateur du plasminogène (t-PA) Humain (Harris *et al.*, 1991) et de la chauve souris, *Desmodus rotundus* (Gohlke *et al.*, 1996). D'autres domaines EGF O-fucosylés ont été décrits sur différents facteurs de coagulation sanguins humains comme le facteur VII (Bjoern *et al.*, 1991), le facteur IX (Nishimura, Takao *et al.*, 1992) et le facteur XII (Harris *et al.*, 1992). Ces trois facteurs de coagulation présentent chacun deux domaines EGF, mais seul le premier porte un O-fucose, le second ne possédant pas le motif consensus. Du fucose O-lié est également présent sur des protéines transmembranaires impliquées dans des processus développementaux majeurs, comme Notch (Moloney, Panin *et al.*, 2000; Moloney, Shair *et al.*, 2000), ses ligands Delta et Serrate/Jagged (Panin *et al.*, 2002) ou encore Cripto (Schiffer *et al.*, 2001; Yan *et al.*, 2002). C'est la comparaison des sites de O-fucosylation de toutes ces protéines qui a permis de déterminer la séquence consensus, $C^2X_{4-5}(S/T)C^3$, nécessaire à cette modification post-traductionnelle. Une recherche dans les banques de données utilisant ce motif comme matrice, permet de trouver de nombreuses protéines potentiellement O-fucosylables comme la fibrilline, l'agrine, la multimérine et le brevican (Moloney et Haltiwanger, 1999).

2.1.1.2 Formation du O-fucosylglycane porté par les EGF

En fonction des domaines EGF, le O-fucose ajouté par Pofut1 peut rester à l'état de monosaccharide ou être allongé en une chaîne oligosaccharidique (Fig.11.). A titre d'exemple, le premier domaine EGF du facteur VII humain porte uniquement un O-fucose alors que celui du facteur IX est substitué par un tétrasaccharide de structure : NeuAc α 2,6Gal β 1,4GlcNAc β 1,3Fuc α 1-S (Harris *et al.*, 1993). De même, les domaines EGF du récepteur Notch et de ses ligands peuvent être modifiés par des O-fucoses seuls, des tétrasaccharides (de structures identiques à celle présente sur le facteur IX à l'exception du type de liaison de l'acide sialique, ici en α 2,3) ou par des formes intermédiaires de l'O-fucosylglycane (Bruckner *et al.*, 2000; Moloney, Panin *et al.*, 2000).

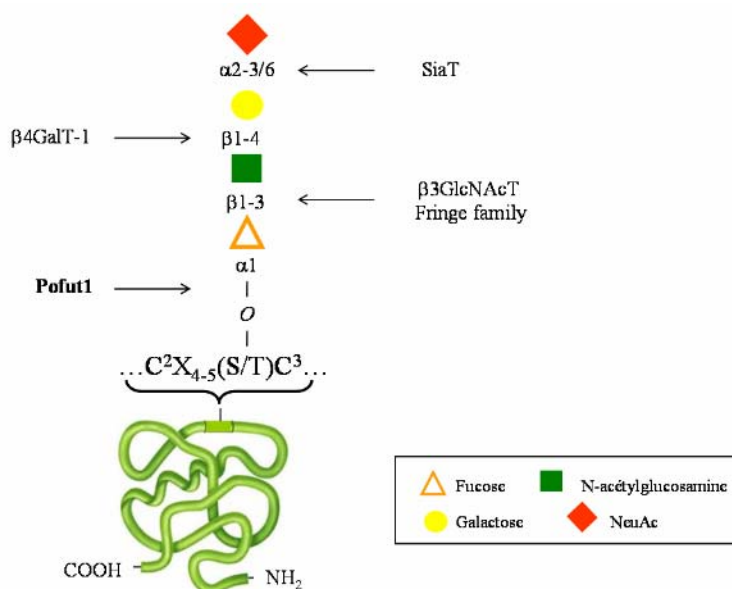


Figure 11 : Structure d'un O-fucosylglycane greffé sur un domaine EGF. (d'après Haltiwanger et Lowe, 2004)

L'allongement du O-fucose en tétrasaccharide commence par l'action d'une β 1,3-N-acétylglucosaminyltransférase de la famille Fringe. La première enzyme de cette famille a tout d'abord été identifiée chez la drosophile (Dfng) comme intervenant dans la détermination de la polarité dorso-ventrale au niveau du cerveau (Irvine et Wieschaus, 1994). Trois enzymes homologues, entre 47 et 64% d'identité en fonction des espèces, ont ensuite été isolées chez les Mammifères : Lunatic Fringe (Lfng), Manic Fringe (Mfng) et Radical Fringe (Rfng) (Cohen *et al.*, 1997; Johnston *et al.*, 1997). Elles semblent être présentes chez l'ensemble des

Gnathostomes (Peterson et McClay, 2005) et possèdent de fortes homologies, à titre d'exemple Rfng de l'Homme et de la souris partagent 83% d'identité (Johnston *et al.*, 1997). Ce sont des enzymes transmembranaires golgiennes qui peuvent également être secrétées grâce à un clivage protéolytique libérant leur domaine catalytique (Yuan *et al.*, 1997). Les enzymes de la famille Fringe transfèrent un résidu GlcNAc depuis un donneur activé, l'UDP-GlcNAc, sur certains résidus fucoses des domaines EGF (Bruckner *et al.*, 2000; Moloney, Panin *et al.*, 2000). Bien qu'homologues, les trois enzymes trouvées chez les Mammifères semblent avoir une spécificité d'action et ne modifient pas forcément les mêmes O-fucoses (Shimizu *et al.*, 2001; Shao *et al.*, 2003). Des tests enzymatiques effectués *in vitro* ont montré que Mfng murine avait une activité plus importante sur le premier domaine EGF O-fucosylé du facteur de coagulation VII humain que ses paralogues Rfng et Lfng; le résultat inverse est observé lorsque l'EGF 26 O-fucosylé de Notch1 humain est utilisé (Rampal, Li *et al.*, 2005). De plus, pour qu'un O-fucose soit allongé, il semble que celui-ci doive se trouver dans un motif peptidique particulier. En effet, Rampal et ses collaborateurs, en 2005, ont identifié les acides aminés E24 et P28 d'un domaine EGF O-fucosylé, comme d'importants indicateurs que celui-ci serait une cible d'action pour les enzymes de la famille Fringe.

L'allongement du O-fucosylglycane porté par un motif EGF se poursuit éventuellement par l'ajout d'un résidu galactose. Cette modification est assurée par une enzyme transmembranaire golgienne, l'enzyme β 4-galactosyltransférase I (β 4Gal-T1). En effet, parmi les six galactosyltransférases présentes dans les cellules CHO, seule cette dernière est capable *in vitro* de modifier le O-fucosylglycane (Chen *et al.*, 2001).

Finalement, le O-fucosylglycane peut être enrichi par un acide sialique sous l'action d' α 2,3- ou α 2,6-sialyltransférases qui sont des enzymes transmembranaires localisées dans le *trans*-Golgi. Chez les Mammifères, elles sont au nombre de 15 (Moloney, Panin *et al.*, 2000). Pour l'heure, celle(s) responsable(s) de l'allongement de l'oligosaccharide n'est (ne sont) pas identifiée(s).

2.1.2 Caractérisation de l'enzyme Pofut1

2.1.2.1 Identification de l'activité *O*-fucosyltransférase 1

Comme nous l'avons vu précédemment, la protéine *O*-fucosyltransférase 1 ou Pofut1 est responsable du greffage du *O*-fucose sur les domaines EGF de certaines protéines possédant le motif consensus de *O*-fucosylation, C²X₄₋₅(S/T)C³. La démarche scientifique qui a amené à la caractérisation de cette enzyme a débuté avec la mise au point en 1996 d'un test *in vitro* de mesure de l'activité *O*-fucosyltransférase 1 (Wang *et al.*, 1996). Le test utilise comme substrat donneur du GDP-*L*-fucose (substrat donneur de toutes les fucosyltransférases), comme substrat accepteur le premier domaine EGF du facteur de coagulation VII humain et un cation divalent. En effet, la présence de ce dernier est nécessaire à l'activité de nombreuses glycosyltransférases (Munro and Freeman, 2000). Dans le cas de Pofut1, plusieurs types de cations divalents ont été testés; c'est avec du manganèse (Mn²⁺) que l'activité enzymatique est la plus forte (Wang *et al.*, 1996).

Ainsi, une première activité *O*-fucosyltransférase 1 a pu être mesurée à partir d'extraits protéiques de cellules CHO et de cellules hépatiques de rat (Wang *et al.*, 1996). Les auteurs ont également montré que le transfert du *O*-fucose n'est possible que sur un domaine EGF complet, c'est-à-dire possédant ses six résidus cystéines nécessaires à l'acquisition de sa structure tridimensionnelle (Fig.10A.).

2.1.2.2 Purification de l'enzyme

Suite à la mise au point du test d'activité, Wang et Spellman (1998) ont purifié l'enzyme Pofut1 à partir de cellules CHO. En se basant sur l'activité enzymatique déjà déterminée et sur l'affinité de l'enzyme pour ses substrats (donneur et accepteur), les auteurs ont utilisé deux étapes de chromatographie d'affinité. La première colonne utilisée contenait l'EGF du facteur VII humain et la deuxième colonne, du GDP-hexanolamine-Sépharose comme substrat donneur. Ils ont ainsi purifié et déterminé les constantes cinétiques de Pofut1 (référencées dans le Tableau 1). L'enzyme *O*-fucosyltransférase 1 représente environ 0,02% des protéines totales des cellules CHO, taux rarement atteint pour une glycosyltransférase.

Tableau 1 : Constantes cinétiques de l'enzyme Pofut1 purifiée à partir de cellules CHO et de l'enzyme Pofut1 humaine.

	Pofut1 des cellules CHO		Pofut1 humaine	
	Km (μM)	Vmax ($\mu\text{mol}/\text{min}/\text{mg}$)	Km (μM)	Vmax ($\mu\text{mol}/\text{min}/\text{mg}$)
EGF-1 du facteur VII (substrat accepteur)	11	2,5	6	3
GDP-fucose (substrat donneur)	5,9-6,4		4	

Après purification de l'enzyme Pofut1, sa glycosylation a pu être mise en évidence après une digestion par la PNGase F (Wang et Spellman, 1998). Cette enzyme coupe les *N*-glycanes oligomannosidiques, hybrides et complexes entre l'asparagine de la chaîne peptidique et la première GlcNAc. La masse moléculaire de Pofut1 passe alors de 44 à 40 KDa, ce qui suggère la présence de *N*-glycanes d'une masse moléculaire totale de 4 KDa. Une digestion avec l'endoglycosidase H, qui coupe entre les deux résidus GlcNAc proximaux, à la fois les *N*-glycanes oligomannosidiques et certains *N*-glycanes hybrides, donne en western blot deux bandes. Elles correspondent à des formes protéiques de masses inférieures à la forme glycosylée. Elles seraient une forme monoglycosylée et une forme déglycosylée de l'enzyme Pofut1. Ainsi, les auteurs concluent que l'enzyme Pofut1 extraite des cellules CHO est une glycoprotéine porteuse de plus d'un oligosaccharide de type oligomannosidique. Toutefois, il ne faut pas écarter l'hypothèse que ces *N*-glycanes puissent être de type hybride.

2.1.2.3 Identification du gène correspondant, *Pofut1*

Suite à la purification et au séquençage peptidique de Pofut1 extraite de cellules CHO, Wang et ses collaborateurs (2001), en s'appuyant sur la séquence N-terminale et sur une recherche d'homologie dans les banques de données, ont pu identifier un EST (Expressed sequence Tag) humain correspondant à cette enzyme. L'ADNc code une protéine de 388 acides aminés et possède deux sites potentiels de *N*-glycosylation (Wang, Shao *et al.*, 2001). Le transcrit de plus de 5 Kb possède une longue partie 3' non traduite, typique de nombreuses glycosyltransférases (Breton et Imberty, 1999), et deux sites potentiels de polyadénylation pouvant expliquer la présence de deux des signaux obtenus en northern blot (Wang, Shao *et al.*, 2001). Cependant, d'autres marquages sur ce même northern blot laisse présager de

l'existence d'autres variants transcriptionnels. Ces différents transcrits semblent être présents dans l'ensemble des tissus humains testés (cerveau, cœur, foie, pancréas, poumons, rein, muscle squelettique, intestin, ovaires, testicules) ce qui accrédi-terait le rôle clé de l'enzyme POFUT1.

Le gène *POFUT1* humain est localisé sur le chromosome 20 au locus q11, près du centromère, entre les gènes *PLAGL2* (Pleiomorphic adenoma gene-like-2) et *KIF3B* (kinesin family member 3B). Une recherche dans les banques de données m'a permis de déterminer qu'il est constitué d'au moins 8 exons. L'épissage alternatif du transcrit primaire de ce gène génère au moins deux transcrits, les variants 1 ([NM_015352](#)) et 2 ([NM_172236](#)) (Fig.12.). Dans le cas du variant 1, c'est l'épissage de l'exon 5 qui est affecté ce qui provoque sa délétion et l'utilisation du second site de polyadénylation dans l'exon 8. Pour le variant 2, seuls les 5 premiers exons sont épissés. Ces deux transcrits codent respectivement deux isoformes protéiques de 388 et 194 acides aminés (Wang, Shao *et al.*, 2001; Luo et Haltiwanger, 2005). Seule la protéine issue de la traduction du variant 1 possède une activité O-fucosyltransférase (Wang, Shao *et al.*, 2001). Les constantes cinétiques sont proches de celles déterminées pour l'enzyme extraite des cellules CHO (Tableau 1).

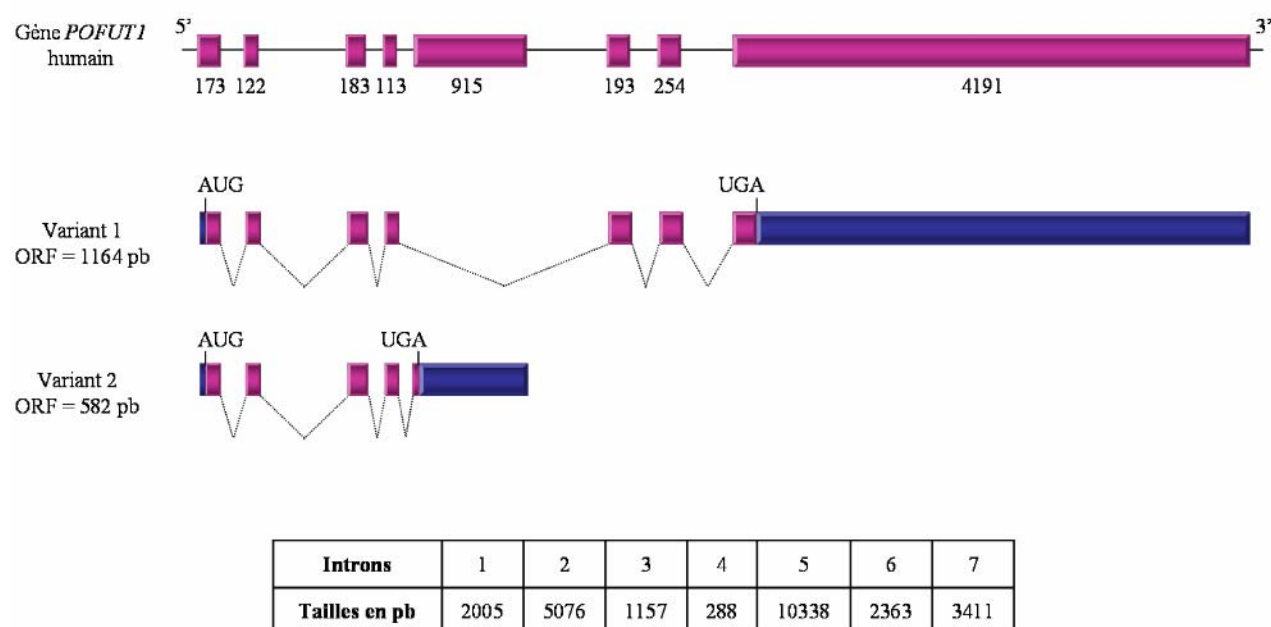


Figure 12 : Représentation de l'organisation génomique du gène *POFUT1* humain et de deux de ses variants transcriptionnels.

Les régions non traduites des transcrits sont en bleu, les tailles des exons sont indiquées sous la structure du gène (en pb), les tailles des introns sont notées dans le tableau ci-dessus.

L'enzyme Pofut1 murine fait 393 acides aminés et conserve 90,4% d'identité avec son homologue humaine (Wang, Shao *et al.*, 2001). Le gène correspondant est constitué d'au moins 7 exons (Shi et Stanley, 2003). Il est localisé sur le chromosome 2 murin au locus H2, dans une région homologue au chromosome 20 humain, puisque retrouvé entre les gènes *Plagl2* et *Kif3b*. Une analyse de son organisation génomique m'a permis de déterminer que les positions des introns étaient strictement conservées par rapport au gène humain et que seul l'équivalent de l'exon 5 était absent. Comme dans les cas de l'Homme, plusieurs transcrits sont détectés par northern blot et présents dans tous les tissus murins analysés : cerveau, cœur, foie, poumons, rate, rein, estomac, intestin, peau, thymus, muscle squelettique et testicules (Shi et Stanley, 2003). Cependant, seul le transcrit majoritaire a été décrit et correspond à la transcription des 7 exons ([NM_080463](#)) (Fig.13.).

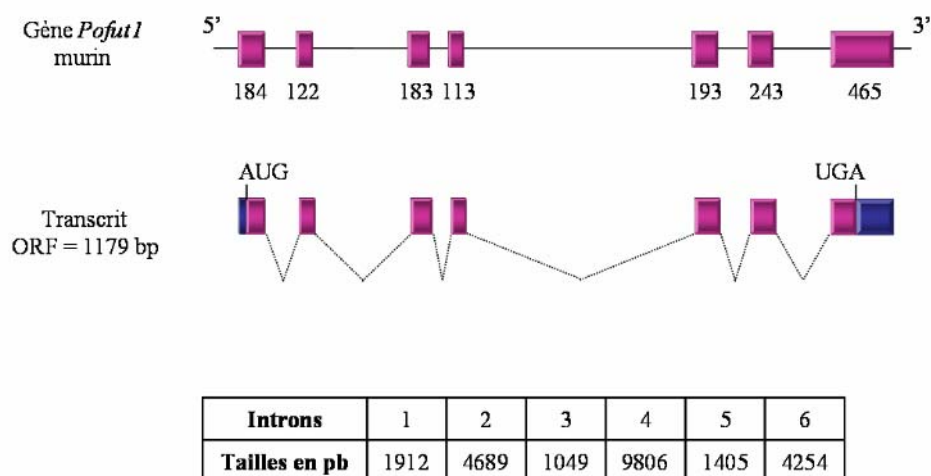


Figure 13 : Représentation de l'organisation génomique du gène *Pofut1* murin et du variant transcriptionnel décrit.

Les régions non traduites du transcrit sont en bleu, les tailles des exons sont indiquées sous la structure du gène (en pb), les tailles des introns sont notées dans le tableau ci-dessus.

L'effort de séquençage systématique des génomes et des transcriptomes permet d'envisager la présence du gène *Pofut1* chez l'ensemble des Bilatériens. Une seule copie du gène est présente dans les génomes entièrement séquencés des représentants de ce groupe, comme par exemple chez *Ciona intestinalis*, *Drosophila melanogaster*, *Rattus norvegicus* et *Takifugu rubripes* (Martinez-Duncker *et al.*, 2003). A l'heure actuelle, hormis les gènes de Mammifères précédemment décrits, seuls les gènes *Pofut1* de *Drosophila melanogaster* et de *Caenorhabditis elegans* ont fait l'objet d'une attention plus particulière (Wang, Shao *et al.*,

2001). Les enzymes correspondantes possèdent respectivement 41,2 et 29,4% d'homologie avec POFUT1 codé par le variant 1 humain.

2.1.2.4 Localisation sub-cellulaire de l'enzyme Pofut1

Lorsqu'en 1996, Wang et ses collaborateurs ont mis au point le test d'une activité O-fucosyltransférase 1, ils ont remarqué que celle-ci était présente dans les fractions solubles et insolubles des extraits protéiques de cellules CHO. Ils en ont déduit que Pofut1, comme de nombreuses glycosyltransférases, devait être une enzyme transmembranaire de type II (Fig.14.), sensible à la protéolyse au niveau de son peptide de liaison (Kitazume *et al.*, 2001).

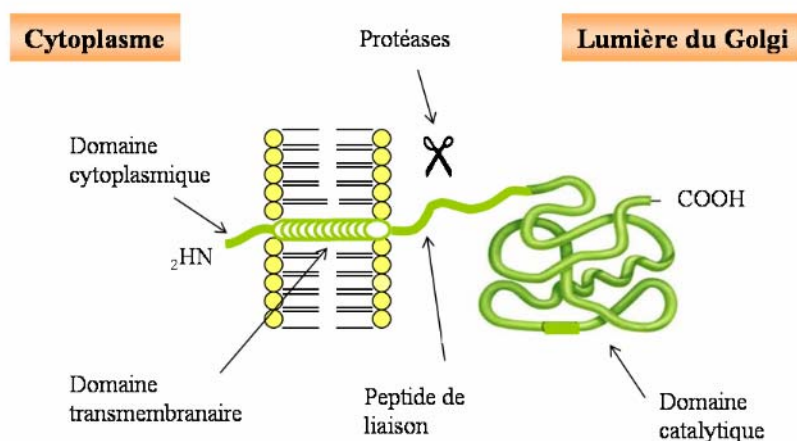


Figure 14 : Représentation schématique d'une glycosyltransférase transmembranaire de type II de l'appareil de Golgi.

Leur hypothèse sur la topologie de l'enzyme Pofut1 fut renforcée en 2001, lors du séquençage de la partie N-terminale de la protéine purifiée, puisque celle-ci était dépourvue des 29 premiers acides aminés (Wang et Spellman, 1998). Ainsi, Pofut1 qui possède des glycanes oligomannosidiques ou hybrides serait une glycoprotéine de type II résidente du réticulum endoplasmique ou du *Cis*-Golgi.

Des études plus récentes (Luo et Haltiwanger, 2005) montrent que l'activité de l'enzyme Pofut1 se retrouve majoritairement dans la fraction soluble d'un extrait protéique de cellules COS-1 (après lyse avec le détergent NP40 et sonication), contrairement à l'enzyme β 4Gal-T1, une galactosyltransférase transmembranaire golgienne prise comme témoin. Le même type d'expérience réalisée après fractionnement de cellules hépatiques de rat confirme

que l'activité O-fucosyltransférase 1 est co-localisée avec l'activité de la Glucose-6-phosphatase, une enzyme réticulaire soluble, et non avec l'activité de l'enzyme β 4Gal-T1. Ces résultats suggèrent fortement que Pofut1 est une enzyme soluble probablement localisée dans la lumière du RE. Pour confirmer ces observations, des expériences d'immunolocalisation ont été conduites sur des cellules COS-1 exprimant la protéine recombinante humaine POFUT1 étiquetée avec 6 histidines (His₆) (Fig.15A. et D.)(Luo et Haltiwanger, 2005). Les auteurs démontrent que POFUT1 est co-localisée avec la calréticuline (Fig.15C.), protéine chaperonne du RE, et non avec GM130 (Fig.15F.), enzyme de l'appareil de Golgi.

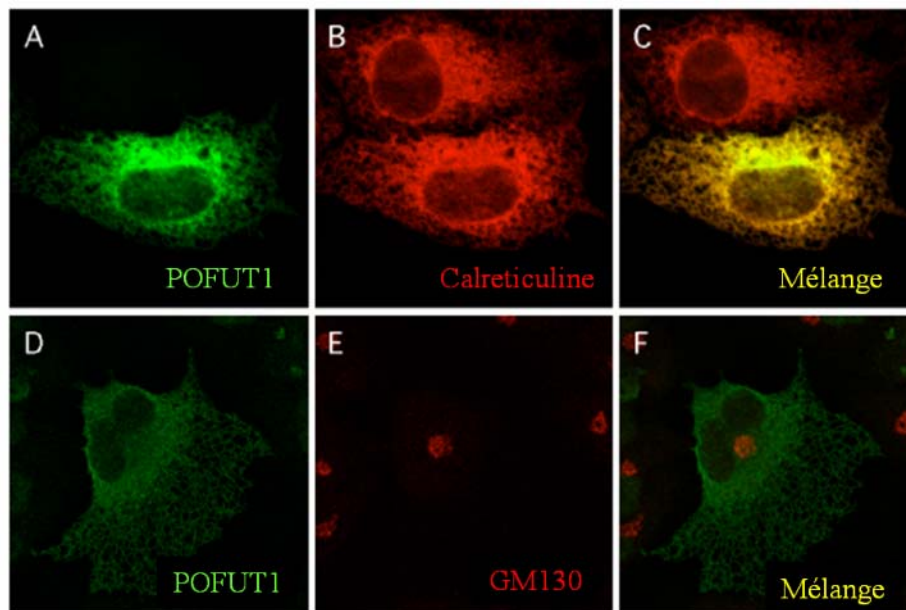


Figure 15 : Immunolocalisation de la protéine recombinante POFUT1 humaine dans les cellules COS-1.

Le marquage de l'enzyme POFUT1 est réalisé à l'aide d'un anticorps anti-His₆. A et D, la protéine POFUT1 est détectée dans le RE; B, la calréticuline est détectée dans le RE; E, GM130 est détecté dans le *Cis*-Golgi; C, superposition des images A et B; F, superposition des images D et E. Extrait de Luo et Haltiwanger, 2005.

Lors de cette même étude, une analyse bioinformatique des séquences protéiques des enzymes Pofut1 humaine, murine et de drosophile, grâce au logiciel SMART (<http://smart.embl-heidelberg.de/>), montre que la partie N-terminale des enzymes ne contient pas de domaine transmembranaire, mais un peptide signal permettant l'entrée des protéines dans le RE (Luo et Haltiwanger, 2005). De plus, la comparaison de ces 3 orthologues révèle la présence d'une séquence KDEL-like (respectivement RDEF, RDEF et HEEL) à l'extrémité C-

terminale des enzymes. Cette courte séquence est un signal de rétention dans le RER (Fig.16.).

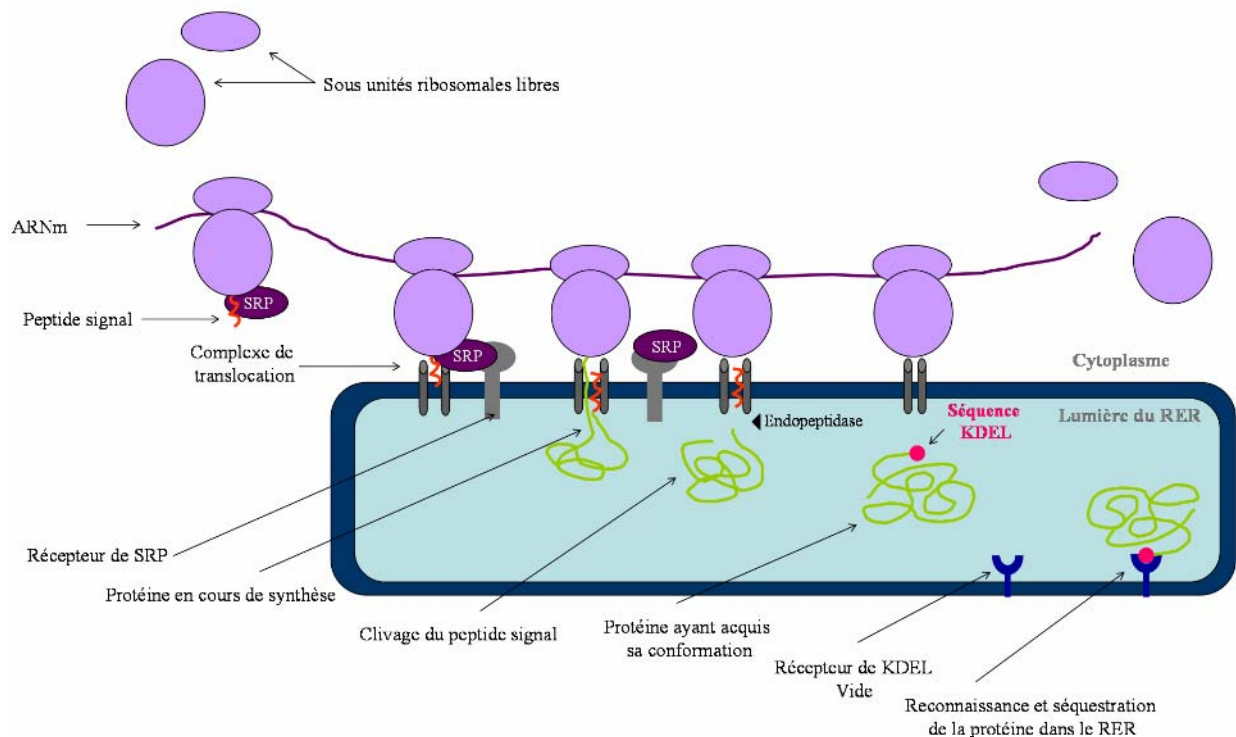


Figure 16 : Représentation schématique de la translocation et de la séquestration de protéines solubles dans le RER.

Dès sa synthèse, le peptide signal est reconnu par une ribonucléoprotéine, la SRP (Signal recognition Particle). L'ensemble (ribosome, ARNm et protéine) est alors adressé à la membrane du RER où SRP se fixe à son récepteur. Au niveau membranaire, le peptide signal est reconnu par le translocon, puis lorsque la synthèse de la protéine est terminée, le peptide est clivé par une endopeptidase. Si la protéine synthétisée présente à son extrémité C-terminale un motif KDEL-like, elle est alors séquestrée dans la lumière du RER.

Le fait que l'enzyme Pofut1 greffe des résidus fucoses uniquement sur des domaines EGF correctement repliés et que celle-ci soit localisée dans le RE, suggère qu'elle pourrait avoir un rôle dans le contrôle du repliement des protéines possédant des domaines EGF. Une étude récente a montré que l'enzyme Ofut1 de drosophile participe au repliement correct du récepteur Notch dans le RE, mais pas à celui de ses ligands (Okajima *et al.*, 2005). En l'absence de cette enzyme, le récepteur Notch n'est pas adressé à la surface des cellules et s'accumule dans le cytoplasme. Ce défaut de localisation peut-être restauré par l'expression d'une protéine Ofut1 ne possédant pas d'activité O-fucosyltransférase, ce qui indique que sa fonction de chaperonne du récepteur Notch est indépendante de son activité enzymatique.

2.1.3 Rôles du O-fucose des domaines EGF

La O-fucosylation, ayant lieu sur des domaines EGF déjà repliés, ne semble pas être directement impliquée dans l'acquisition de leur structure tridimensionnelle. A titre d'exemple, la O-fucosylation du facteur de coagulation VII humain ne semble modifier ni sa structure ni sa fixation aux facteurs tissulaires (Kao *et al.*, 1999). Cependant, celle-ci pourrait avoir un rôle dans la stabilisation des sites de fixation du Ca^{2+} que possèdent de nombreux domaines EGF. Ainsi, le calcium permettrait une augmentation de l'affinité des complexes ligands/récepteurs (Kao *et al.*, 1999).

La défucosylation du domaine EGF de l'activateur du plasminogène réalisée par semi-hydrolyse à l'acide trifluoroacétique n'empêche pas sa fixation à son récepteur mais abolit son activité mitogène. Les mécanismes mis en jeu ne sont pas encore déterminés mais l'absence de O-fucose semble empêcher l'activation de la cascade de signalisation intra-cytoplasmique (Rabbani *et al.*, 1992).

Le rôle du O-fucose lié à un domaine EGF a également été montré, *in vitro*, dans le cas de la signalisation induite par Cripto. Cette protéine liée à la membrane de certaines cellules par une ancre GPI (Glycosylphosphatidylinositol) est impliquée dans l'établissement de la polarité antéro-postérieure durant l'embryogenèse (Ding *et al.*, 1998). La mutation du site de O-fucosylation présent dans l'unique domaine EGF-like de Cripto murin (Fig.17.) réduit son affinité pour son cofacteur Nodal, membre de la super famille des $\text{TGF}\beta$, et donc sa capacité à induire une cascade de signalisation intra-cytoplasmique conduisant à l'expression de gènes spécifiques du mésoderme (Schiffer *et al.*, 2001).

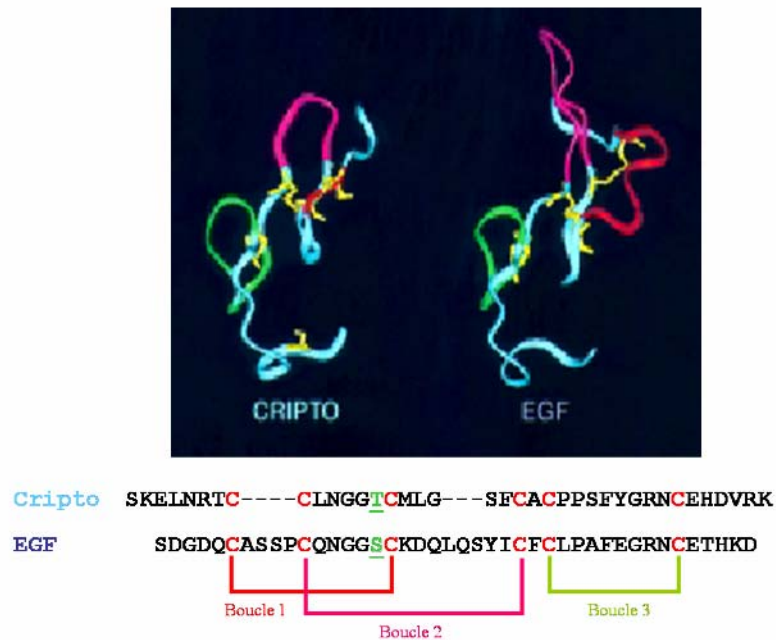


Figure 17 : Modélisation du domaine EGF-like de Cripto murin et du premier domaine EGF du facteur de coagulation VII Humain.

Les six cystéines participant à la formation des ponts disulfures sont indiquées en rouge. Le site de O-fucosylation (S ou T) est en vert et souligné. Remarquez que les boucles 1 et 2 des domaines EGF-like sont plus petites que celles des domaines EGF. D'après Shiffer *et al.*, 2001.

> Exemple des récepteurs Notch et de leurs ligands

La déficience en fucose d'une lignée cellulaire CHO (Lec 13, mutante pour une des étapes de la production de GDP-fucose) induit des perturbations dans l'initiation de la cascade de signalisation dépendante des récepteurs Notch (Chen *et al.*, 2001). Il semblerait également que la O-fucosylation des récepteurs Notch permette de réduire la *cis*-activation de ceux-ci par leurs ligands, en particulier lors de leur transit dans l'appareil de Golgi avant leur sécrétion (Sakamoto *et al.*, 2002).

Structure moléculaire des récepteurs Notch et de leurs ligands

La majeure partie des données disponibles actuellement concernant la O-fucosylation par l'enzyme Pofut1 a été obtenue grâce aux divers travaux réalisés sur les membres de la famille des récepteurs Notch et sur leurs ligands. En interagissant, ils régulent de nombreuses voies de signalisation notamment durant le développement embryonnaire. Ces mécanismes de

régulation étant extrêmement complexes, il est avant tout nécessaire de décrire la structure de ces protéines cibles de la O-fucosylation.

> Au niveau moléculaire, le récepteur Notch est une protéine transmembranaire de type I (Fig.18.), impliquée dans de nombreux processus développementaux comme la neurogenèse, la somitogenèse, la cardiogenèse, la vasculogenèse, etc. Les récepteurs Notch, au nombre de 4 chez les Mammifères (Notch 1 à 4) contre 1 chez la drosophile (dNotch), sont bien conservés dans leurs structures. Ils sont composés d'un fragment extracellulaire (ECN) et d'un fragment transmembranaire (NTM) (Fig.18.). Le fragment extracellulaire est constitué de nombreux domaines EGF (36 EGF pour dNotch et Notch 1 et 2; 34 et 24 EGF respectivement pour Notch 3 et 4) et de motifs LNR (Lin12/Notch repeats). Les LNR semblent restreindre l'activation non spécifique (ligand-indépendant) du récepteur et notamment sa protéolyse par des métalloprotéases (Sanchez-Irizarry *et al.*, 2004). Le fragment NTM est composé d'un court domaine extracellulaire, d'un domaine transmembranaire unique et d'un domaine intracellulaire (ICN). L'ICN présente un domaine RAM (Regulation of Amino-acid Metabolism) régulateur de la transcription des gènes cibles, deux domaines de localisation nucléaire (NLS), des domaines ankyrine répétés (ANK) impliqués dans les interactions protéine/protéine (Wharton *et al.*, 1985; Blaumueller et Artavanis-Tsakonas, 1997) et une région PEST (Pro-Glu-Ser-Thr) qui pourrait intervenir dans le processus de dégradation de l'ICN par ubiquitination (Greenwald, 1998; Artavanis-Tsakonas *et al.*, 1999). Les récepteurs Notch1, 2 et dNotch possèdent également une région de réponse aux cytokines (NCR) ainsi qu'un domaine TAD (transcriptional activation domain). Notch 3 est dépourvu de ce dernier domaine.

Le récepteur Notch est synthétisé sous la forme d'un précurseur de 300 KDa qui est clivé dans l'appareil de Golgi par une furine (Blaumueller et Artavanis-Tsakonas, 1997; Logeat *et al.*, 1998; Guidos, 2002). Ce clivage conduit à la formation des deux sous-unités, ECN et NTM, qui restent associées de manière non-covalente et calcium-dépendante (Annaert et De Strooper, 1999). L'hétérodimère est ensuite adressé à la surface cellulaire où il reste inactif, jusqu'à ce qu'il lie un de ses ligands présents à la surface d'une cellule contiguë.

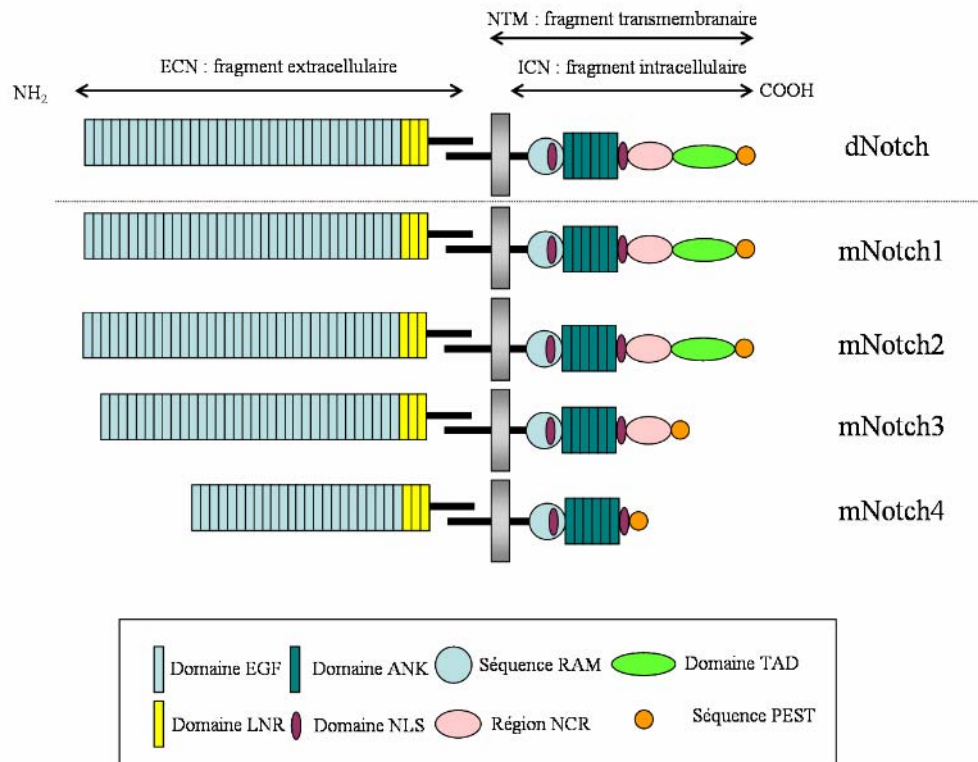


Figure 18 : Domaines conservés des récepteurs Notch chez la drosophile (dNotch) et les Mammifères (mNotch1 à 4).

> Les structures des ligands interagissant avec les récepteurs Notch sont également bien conservées entre les espèces animales. Ce sont aussi des protéines transmembranaires de type I. Chez la drosophile (Fig.19.), les ligands de Notch sont codés par les gènes *Delta* et *Serrate* (Doherty *et al.*, 1996). La partie extracellulaire du ligand Delta contient neuf domaines EGF et un domaine DSL, nommé ainsi sur la base des noms des ligands de Notch, à savoir Delta, Serrate et Lag-2 chez *Caenorhabditis elegans*. Chez les Mammifères, trois gènes homologues à *Delta* ont été décrits (Fleming, 1998). Ils codent les protéines Delta-like 1, 2 et 3 (Dl1, Dl2 et Dl3) (Fig.19.). Elles contiennent chacune un domaine DSL et respectivement, 8, 6 et 8 domaines EGF. Le deuxième ligand de Notch est Serrate chez la drosophile et Jagged chez les Mammifères où il existe 2 paralogues, appelés Jagged1 et 2 (Fleming, 1998; Lissemore et Starmer, 1999). Leur fragment extracellulaire est composé d'un domaine DSL, de 14 domaines EGF pour Serrate et 16 pour Jagged1 et 2, ainsi que d'un troisième domaine impliqué probablement dans les processus de dimérisation (VWC). Ces ligands possèdent un domaine riche en cystéines dans la partie N-terminale, jouant potentiellement un rôle dans la spécificité des liaisons ligand/récepteur. De nouveaux ligands de Notch ont été décrits, tels que la F3/Contactine et la NB3 qui activent la voie Notch

et jouent un rôle dans la différenciation des oligodendrocytes (Hu *et al.*, 2003; Cui *et al.*, 2004). Ils diffèrent des ligands cités précédemment par l'absence de motifs EGF et la présence de nombreux domaines immunoglobulines. Cette multiplication des acteurs impliqués dans la voie de signalisation par Notch génère une complexité croissante, de telle sorte qu'à l'heure actuelle, son fonctionnement reste pour partie mal connu.

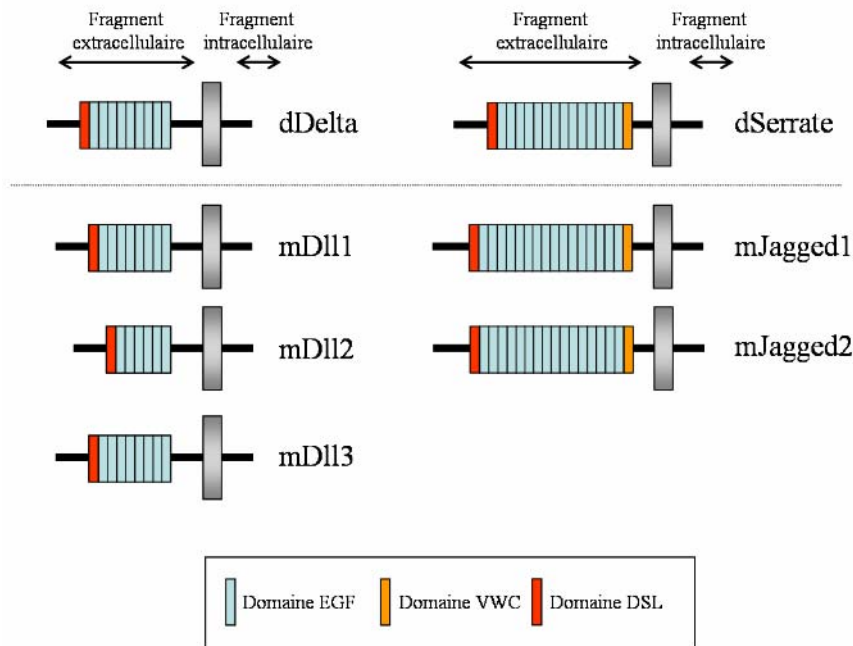


Figure 19 : Domaines conservés des ligands de Notch chez la drosophile et les Mammifères.

La voie de signalisation par Notch définit un mécanisme d'interactions cellulaires qui s'exerce essentiellement entre cellules voisines et qui contrôle de nombreux processus développementaux parmi lesquels le choix de l'identité cellulaire, la prolifération, la différenciation et l'apoptose. La voie de signalisation par Notch est retrouvée chez tous les organismes pluricellulaires étudiés jusqu'à présent, des nématodes à l'Homme (Weinmaster, 1997; Kimble et Simpson, 1997; Artavanis-Tsakonas *et al.*, 1999; Mumm et Kopan, 2000; Baron, 2003). Nous reviendrons sur les cascades de signalisation déclenchées par les récepteurs Notch dans le chapitre 3 de ce manuscrit.

Rôle du O-fucose présent sur les récepteurs Notch et leurs ligands

De nombreux travaux soulignent que les O-fucoses présents sur les récepteurs Notch et leurs ligands jouent un rôle fondamental dans leurs interactions. Chez la drosophile, l'implication de la O-fucosylation des domaines EGF a été examinée après surexpression du gène *Ofut1* ou réduction de son expression par ARN interférence (Okajima et Irvine, 2002), et plus récemment, par l'isolement d'un mutant *Ofut1* *-/-* nommé *neurotic* (Sasamura *et al.*, 2003). Dans tous les cas, le phénotype observé est proche de celui d'un mutant pour le gène *Notch* avec comme conséquences des malformations au niveau de l'abdomen, des pattes, des ailes, du cerveau et des yeux. Le O-fucose serait donc nécessaire au fonctionnement correct du récepteur Notch. Ainsi, 23 des 36 domaines EGF du récepteur dNotch, 7 des 9 domaines EGF de Delta et 8 des 14 domaines EGF de Serrate sont potentiellement O-fucosylables. La présence et les fonctions respectives des O-fucoses sur les domaines EGF commencent à être étudiées. Chez la drosophile, il semble que les O-fucoses portés par les domaines EGF 11 et 12 de dNotch soient primordiaux pour son interaction avec ses ligands (Lei *et al.*, 2003). Une mutation du site de O-fucosylation sur le domaine EGF 12 de la protéine Notch1 murine entraîne une inhibition de son interaction avec les ligands Delta1 et Jagged1. A la différence, une mutation du site de O-fucosylation sur l'EGF 26 entraîne une hyper-activation de Notch1 par Delta1 et Jagged1. Une localisation incorrecte du récepteur Notch est même observée dans le cas d'une mutation sur l'EGF 27 (Rampal, Arbolaveda-Velasquez *et al.*, 2005). Bien que les mécanismes de régulation des récepteurs Notch soient plus complexes chez les Mammifères que chez la drosophile, ceci étant dû à l'implication d'un plus grand nombre de partenaires protéiques, le O-fucose semble y jouer un rôle similaire. En effet, comme nous le verrons plus loin, une souris KO pour le gène *Pofut1* meurt à l'état embryonnaire suite à de nombreux dysfonctionnements dans divers processus développementaux, comme la somitogenèse et la neurogenèse. Le phénotype obtenu est, en fait, le même que pour un KO de gènes dont les produits interviennent dans les cascades de signalisation intracellulaire initiées par les quatre récepteurs Notch (Shi et Stanley, 2003). Ceci confirme, s'il le fallait, le rôle primordial de la O-fucosylation dans les voies de signalisation par Notch et par conséquent dans le développement embryonnaire.

Le greffage de résidus O-fucoses par l'enzyme Pofut1 sur les domaines EGF de Notch et de ses ligands est un événement incontournable dans le déclenchement des voies de signalisation, puisque c'est leur allongement ou non par les enzymes de la famille Fringe qui

va moduler leurs interactions. En effet, l'action des enzymes de la famille Fringe sur certains domaines EGF du récepteur Notch de drosophile, les EGF 12 et 24 à 28, va potentialiser son interaction avec Delta et inhiber celle avec Serrate (Rampal, Arbolaveda-Velasquez *et al.*, 2005). Chez la drosophile, un KO du gène *Fringe* peut être complété par n'importe laquelle des enzymes Fringe de souris (Johnston *et al.*, 1997). Cependant, les fonctions de Notch se trouvent être modifiées dans certains organes comme le cerveau et au niveau de certaines structures comme les pattes (Rampal, Li *et al.*, 2005). Une souris KO pour le gène *Lfng* présente des défauts dans la formation des muscles squelettiques suite à un problème de somitogenèse et meurt peu après la naissance, alors qu'une souris KO pour le gène *Rfng* ne présente aucun phénotype particulier. Le phénotype du double mutant *Lfng/Rfng* *-/-* est le même que celui de *Lfng* *-/-* (Zhang, Norton *et al.*, 2002). Pour l'instant, aucune donnée de ce type n'est disponible concernant le gène *Mfng*. L'allongement du O-fucosylglycane par l'enzyme β 4Gal-T1 semble également nécessaire à la modulation de l'interaction entre Notch et ses ligands. En effet, bien qu'une souris mutante pour ce gène naisse sans problème apparent, l'observation de son développement embryonnaire montre de sévères altérations de la somitogenèse (Chen *et al.*, 2006). Le phénotype discret de l'adulte mutant est sûrement dû à une activité redondante des β 4Gal-Ts. L'allongement en tétrasaccharide ne semble pas absolument nécessaire à la signalisation par Notch et ses ligands. En effet, les cellules CHO lec2 déficientes dans le transport du CPM-acide sialique, du fait de l'expression d'isoformes protéiques inactives du transporteur présent dans la membrane de l'appareil de Golgi, présentent une voie de signalisation par Notch normale (Chen *et al.*, 2001).

2.2 Découverte récente d'une enzyme homologue à Pofut1, la *O*-fucosyltransférase 2

2.2.1 Une seconde cible de la *O*-fucosylation, les TSR

2.2.1.1 Protéines possédant des TSR *O*-fucosylés

Comme nous le verrons par la suite, la caractérisation de l'activité de l'enzyme Pofut2 est récente. Nous savons aujourd'hui, qu'elle est responsable du greffage de résidus fucoses sur les acides aminés sérine et thréonine présents dans les motifs TSR (Thrombospondin type I repeats). Les premiers domaines TSR *O*-fucosylés décrits ont été ceux de la trombospondine I humaine (Hofsteenge *et al.*, 2001; Gonzalez de Peredo *et al.*, 2002). C'est une glycoprotéine présente sur les granules α des plaquettes sanguines qui a un rôle dans leur agrégation (Legrand *et al.*, 1992) et dans le maintien de la structure des filaments de fibrine lors de la formation des caillots sanguins (Bale, 1987). Elle peut également être sécrétée par des cellules en croissance et interagir avec des protéines de surface, de la matrice extracellulaire ou des protéines solubles. De par ses rôles dans la modulation des interactions cellulaires, elle participe à de nombreux processus physiologiques : développement du système nerveux, immunité, angiogenèse, inflammation, activation des facteurs de croissance (Mosher, 1990; Frazier, 1991; Roberts, 1996; Adams et Tucker, 2000).

La trombospondine I humaine possède 3 domaines TSR, tous *O*-fucosylés. D'autres protéines de la même famille telles que la properdine humaine (un régulateur positif de l'activation du complément) et la F-spondine de rat (protéine de la matrice extracellulaire qui interviendrait entre autres dans le développement de la moëlle épinière et du système nerveux) comportent respectivement 6 et 5 domaines TSR dont 4 sont *O*-fucosylés (Hofsteenge *et al.*, 2001; Gonzalez de Peredo *et al.*, 2002). C'est la comparaison des sites de *O*-fucosylation de

ces protéines qui a permis de définir le motif peptidique consensus, $WX_5C^1X_{2-3}(S/T)C^2X_2G$, nécessaire à la O-fucosylation des TSR (Hofsteenge *et al.*, 2001). Une analyse *in silico* des domaines TSR présents dans 88 protéines humaines indique que 63 d'entre-elles (la thrombospondine 2, certaines sémaphorines, divers facteurs du complément, etc.) possèdent ce motif consensus, suggérant que la O-fucosylation des domaines TSR serait répandue.

2.2.1.2 Formation du O-fucosylglycane porté par les TSR

Le O-fucose présent sur les domaines TSR peut rester à l'état de monosaccharide. Néanmoins, la plupart du temps, celui-ci est allongé en disaccharide de structure : $Glc\beta 1,3Fuc\alpha 1-S$ (Hofsteenge *et al.*, 2001; Gonzalez de Peredo *et al.*, 2002). En effet, les 3 domaines TSR de la thrombospondine I humaine et 4 de la properdine sont modifiés par cette structure disaccharidique. Cependant, à l'heure actuelle, aucune information n'est disponible à propos de l'allongement éventuel des O-fucoses présents sur la F-spondine (Gonzalez de Peredo *et al.*, 2002). Ce O-fucosylglycane que l'on sait désormais porté par les domaines TSR avait été depuis longtemps mis en évidence par chromatographie, dans l'urine humaine (Hallgren *et al.*, 1975), puis plus récemment, dans les cellules CHO (Moloney *et al.*, 1997), sans pour autant savoir sur quelles protéines il était greffé.

En 1999, Moloney et Haltiwanger caractérisent une activité fucose- $\beta 1,3$ -glucosyltransférase ($\beta 3$ -GlcT) dans les cellules CHO. Ils utilisent alors un accepteur qui est un composé aromatique de faible poids moléculaire, le para-Nitrophényl- α -L-fucose (pNp- α -L-fucose), substrat qui s'affranchit de la portion peptidique à l'époque non connue. La $\beta 3$ -GlcT contenue dans la fraction soluble est capable de transférer le glucose d'un substrat donneur activé, l'UDP-glucose, sur le résidu fucose, ceci en présence d'ions Mn^{2+} . Une telle activité a aussi pu être mesurée dans des extraits protéiques cellulaires chez différentes espèces comme l'Homme, la souris, le rat et le poulet (Moloney et Haltiwanger 1999). Récemment, il a été démontré que c'est cette enzyme qui est responsable de l'allongement en disaccharide du O-fucose présent sur les domaines TSR (Luo, Koles *et al.*, 2006; Luo, Nita-Lazar *et al.*, 2006) (Fig.20.). Toutefois, nous ne disposons d'aucune information au niveau moléculaire concernant le gène ou la protéine assurant cette fonction enzymatique.

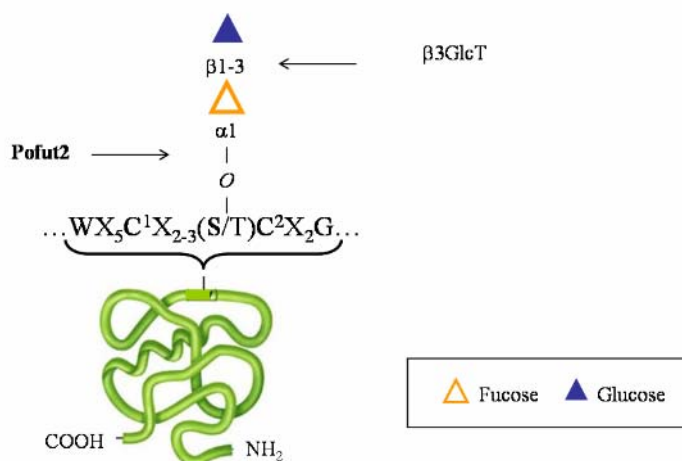


Figure 20 : Structure d'un O-fucosylglycane greffé sur un domaine TSR.
d'après Luo, Nita-Lazar *et al.*, 2006.

2.2.2 Caractérisation de l'enzyme Pofut2

Pofut2 fut citée pour la première fois dans la littérature, en 2002 par Roos et ses collaborateurs, lors d'une étude globale du génome de la drosophile qui concernait l'identification de gènes codant des enzymes impliquées dans le métabolisme des glycanes fucosylés. Cette étude leur a permis de reconnaître, sur la base d'homologies avec Pofut1, ce second gène codant potentiellement une O-fucosyltransférase chez la drosophile et l'Homme, nommé respectivement *Ofut2* et *POFUT2* (Roos *et al.*, 2002). Une analyse phylogénétique montre que ce gène est présent chez de nombreux autres Métazoaires et chez certains protozoaires du genre *Plasmodium* (Martinez-Duncker *et al.*, 2003).

2.2.2.1 Identification de l'activité O-fucosyltransférase 2

Ce n'est que cette année qu'un test d'activité a mis en évidence la présence d'une enzyme, dans les cellules CHO, capable de O-fucosyler les TSR (Luo, Panin *et al.*, 2006). Pour sa réalisation, du GDP-L-fucose (substrat donneur), le troisième domaine TSR de la thrombospondine I humaine (substrat accepteur) et du Mn^{2+} (cation divalent) ont été utilisés. Dans le même temps, les auteurs ont aussi démontré que l'enzyme *Ofut2* de drosophile est capable de réaliser ce greffage *in vitro*. Toutefois, un domaine TSR possédant une conformation tridimensionnelle correcte impliquant la formation de trois ponts disulfures est indispensable à la catalyse enzymatique. De façon plus étonnante, l'enzyme *Ofut2*, à la

différence de la majorité des fucosyltransférases, est toujours active en absence de cation divalent (Luo, Koles *et al.*, 2006). L'activité enzymatique de Pofut2 sur des domaines TSR serait cependant beaucoup plus faible que celle mise en évidence pour l'enzyme Pofut1 sur des domaines EGF. De façon légitime, on peut donc s'interroger sur l'existence d'autres substrats accepteurs vis à vis desquels l'activité de Pofut2 serait plus forte.

2.2.2.2 Etablissement de la structure du gène *Pofut2*

C'est avant de découvrir les fonctions de l'enzyme Pofut2, en étudiant les gènes du chromosome 21 humain pouvant être impliqués dans le phénotype du syndrome de Down, que Menzel et ses collaborateurs en 2004 ont identifié et cartographié le gène *POFUT2* humain au locus 21q22.3. Ce gène est constitué d'au moins 10 exons et l'épissage alternatif de son transcrite primaire donne naissance à 3 variants transcriptionnels. Le variant A ([NM_015227](#)) est issu du non épissage des introns 8 et 9, le variant B ([NM_133634](#)), d'un épissage alternatif dans l'exon 8 et le variant C ([NM_133635](#)), du même épissage dans l'exon 8, ce variant ne possédant pas l'exon 9. Ils codent respectivement des protéines de 424, 380 et 429 acides aminés (Fig.21.). Parmi ces trois variants, seul le C coderait la protéine potentiellement active chez l'Homme (Luo, Nita-Lazar *et al.*, 2006).

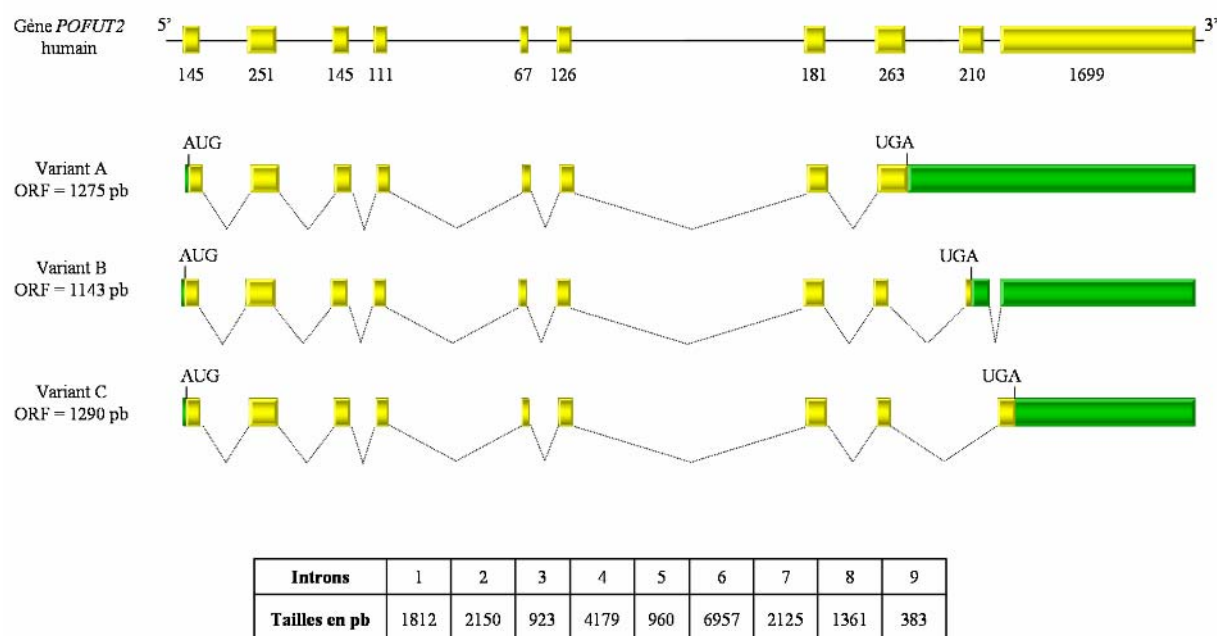


Figure 21 : Représentation de l'organisation génomique du gène *POFUT2* humain et de ses variants transcriptionnels.

Les régions non traduites des transcrits sont en vert; les tailles des exons sont indiquées sous la structure du gène (en pb); les tailles des introns sont notées dans le tableau ci-dessus.

Menzel et ses collaborateurs (2004) ont également identifié *in silico*, les gènes *Pofut2* du nématode *Caenorhabditis elegans* et de la souris. Les protéines correspondantes possèdent respectivement 45% ([Q8WR51](#)) et 92% ([NP_084538](#)) d'identité avec la protéine issue du variant C humain. Chez *C. elegans*, le gène *Pofut2* (aussi nommé *Pad-2* pour patterning defective-2), situé sur le chromosome III, est constitué de 5 exons répartis sur environ 2Kb. Chez la souris, il se trouve sur le chromosome 10 au locus C1 et comporte au moins 9 exons. Pour l'instant, un seul transcrit murin, codant une protéine de 429 acides aminés et orthologue du variant C humain, est répertorié dans les banques de données (Fig.22.).

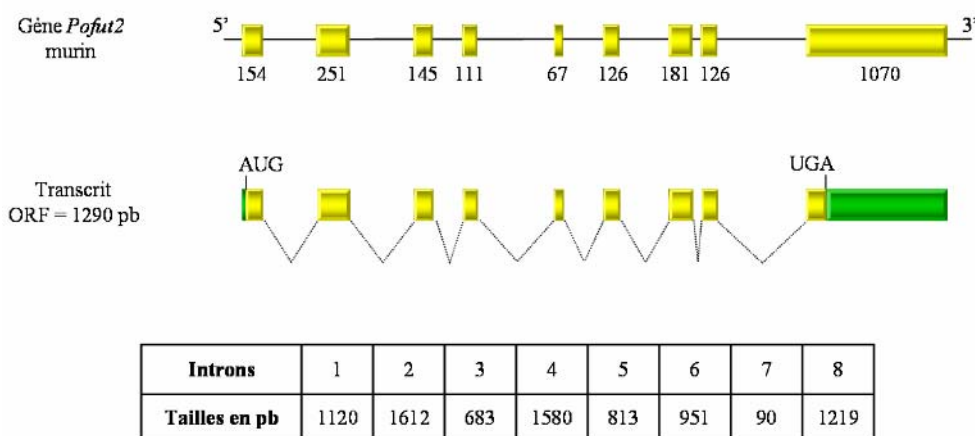


Figure 22 : Représentation de l'organisation génomique du gène *Pofut2* murin et de son variant transcriptionnel.

Les régions non traduites des transcrits sont en vert; les tailles des exons sont indiquées sous la structure du gène (en pb); les tailles des introns sont notées dans le tableau ci-dessus.

Le gène *Pofut2* est présent en une seule copie dans le génome de tous les Métazoaires étudiés jusqu'à présent (Menzel *et al.*, 2004). Il est relativement bien conservé. Tout comme pour le gène *POFUT1*, des transcrits, correspondant au gène *POFUT2* sont détectés dans tous les tissus humains analysés (colon, foie, rate, reins, cœur et cerveau) et correspondent majoritairement aux variants B et C (Menzel *et al.*, 2004). Les données bibliographiques à propos de *Pofut2* sont peu nombreuses et la plupart très récentes. Néanmoins, sa large répartition tissulaire laisse présager d'un rôle fondamental pour la O-fucosylation de type 2 dans le métabolisme cellulaire.

2.2.2.3 Localisation sub-cellulaire de l'enzyme Pofut2

La localisation exacte de l'enzyme Pofut2 reste sujette à discussion. La transfection de cellules COS-7 avec une construction permettant l'expression d'une protéine POFUT2 humaine fusionnée avec un marqueur fluorescent (eGFP) révèle une localisation golgienne (Menzel *et al.*, 2004). Cependant, ces expériences ont été conduites avec l'ORF du variant B humain qui code potentiellement une protéine inactive due à un épissage alternatif du transcrit et à l'usage d'un codon stop prématuré. Dernièrement, une localisation réticulaire pour l'enzyme Pofut2 a été démontrée (Luo, Nita-Lazar *et al.*, 2006). Des expériences d'ultracentrifugation sur des extraits protéiques de cellules COS-1 révèlent une activité *O*-fucosyltransférase 2 dans le surnageant, contrairement à l'activité β 4-galactosyltransférase 1 (témoin de localisation golgienne) qui se trouve dans le culot. Pofut2 serait donc une protéine soluble. L'analyse de l'extrémité N-terminale des séquences peptidiques de Pofut2 chez l'Homme, la souris, la drosophile et *Caenorhabditis*, montre l'existence d'un peptide signal d'adressage au RE. Néanmoins, celles-ci ne possèdent pas de signal de rétention du RE de type KDEL. Toutefois, il est possible que d'autres motifs non consensus soient impliqués dans le mécanisme de rétention réticulaire.

Les *O*-fucosyltransférases 2 de Mammifères possèdent 3 sites potentiels de *N*-glycosylation. Aucune étude n'a encore porté sur ceux-ci. Cependant, si l'enzyme Pofut2, tout comme Pofut1, est effectivement localisée dans le RE, on peut penser que les *N*-glycanes sont de type oligomannosidique voire hybride.

2.2.3 Rôles du *O*-fucose des domaines TSR

La thrombospondine I (TSP1) est une protéine *O*-fucosylée impliquée dans de nombreux processus biologiques (agrégation des plaquettes sanguines, maintien de la structure des filaments de fibrine, interactions cellulaires). Bien que le rôle des *O*-fucosylglycanes qu'elle porte ne soit pas déterminé pour l'instant, il semblerait que les régions peptidiques *O*-fucosylées soient impliquées dans l'interaction de la protéine avec de nombreux partenaires (Adams et Tucker, 2000). Par exemple, son activité anti-angiogénique serait due à l'interaction de la séquence YC¹SVTC²G du TSR1 avec le CD36 présent à la surface des cellules endothéliales (Asch *et al.*, 1992). De même, la séquence C¹SVTC²G présente sur les TSR2 et 3 de TSP1 constituerait un site important pour les interactions

cellule-cellule (Asch *et al.*, 1992). La O-fucosylation des TSR pourrait donc influencer les interactions entre des protéines extracellulaires de façon analogue aux O-fucosylglycanes présents sur les récepteurs Notch.

La seule étude réalisée *in vivo* afin d'appréhender le rôle de la O-fucosylation de type 2 concerne le nématode *Caenorhabditis elegans* (Menzel *et al.*, 2004). Elle montre que l'inhibition de l'expression du gène *Pofut2* par ARN interférence entraîne la mort de 17% des embryons. Sa surexpression semble avoir une pénétrance plus forte puisque 83% des embryons meurent. Dans les deux cas, les individus survivants présentent d'importants problèmes développementaux avec arrêt du développement dans certains cas, mauvais positionnement des organes (muscles, intestin), infertilité, problèmes neuronaux, malformations du corps et difficulté de mouvements. Comme *Pofut1*, *Pofut2* serait un élément incontournable du bon développement embryonnaire et de la survie d'un organisme.

2.3 Caractéristiques moléculaires et fonctionnelles des O-fucosyltransférases

Les deux enzymes Pofut1 et Pofut2 ont chacune une activité bien définie. L'enzyme Pofut1 ne greffe pas de fucose sur les TSR. De même, Pofut2 n'utilise pas les EGF comme accepteurs. De plus, les glycosyltransférases responsables de l'élongation des O-fucoses sont différentes selon l'accepteur (EGF ou TSR). En effet, les membres de la famille Fringe sont incapables d'allonger le O-fucose présent sur un TSR. Enfin, il n'est pas connu de situation où le O-fucose greffé sur un EGF est allongé par un glucose. Ces observations prouvent que l'action des enzymes Fringe et β 3-GlcT s'accomplit par la reconnaissance simultanée du O-fucose et du motif peptique (EGF ou TSR) qui le porte (Luo, Koles *et al.*, 2006).

Le fait que Pofut2, tout comme Pofut1, ne fucosyle que des accepteurs peptidiques correctement repliés, suggère un rôle de ces enzymes dans le «contrôle qualité» du repliement des protéines possédant des domaines EGF ou TSR et présentes dans la lumière du RE. Cependant, le fait que très probablement ces deux enzymes soient réticulaires soulève un problème d'accessibilité à leur substrat donneur, le GDP-L-fucose. En effet, ce dernier est synthétisé dans le cytoplasme (Fig.23.) et jusqu'à présent, seul un transporteur golgien est connu (Abeijon *et al.*, 1997). Une perte de fonction de ce transporteur chez l'Homme, due à des mutations génétiques, est à l'origine d'un syndrome de type CDG, nommé LADII (Leukocyte Adhesion Deficiency type II), caractérisé notamment par un retard de croissance, des problèmes immunitaires et neurologiques (Etzioni *et al.*, 1992). Des analyses effectuées sur les protéines de patients atteints de ce syndrome montrent que la fucosylation des N-glycanes est extrêmement réduite, alors que la O-fucosylation au niveau des EGF et TSR est beaucoup moins affectée (Lubke *et al.*, 1999 et 2001; Sturla *et al.*, 2001; Etzioni *et al.*, 2002). De plus, une étude récente sur des mutants de drosophile dont le gène codant le transporteur de GDP-fucose (*Gfr*) a été invalidé, montre que la O-fucosylation du récepteur Notch est affectée, mais à un degré moindre que la fucosylation des N-glycanes qu'il porte (Ishikawa *et al.*, 2005). Quoi qu'il en soit, les mutants homozygotes sont viables. Ces résultats suggèrent

l'existence d'un autre mécanisme, conservé chez la drosophile et les Mammifères, de transport du GDP-fucose dans le RE.

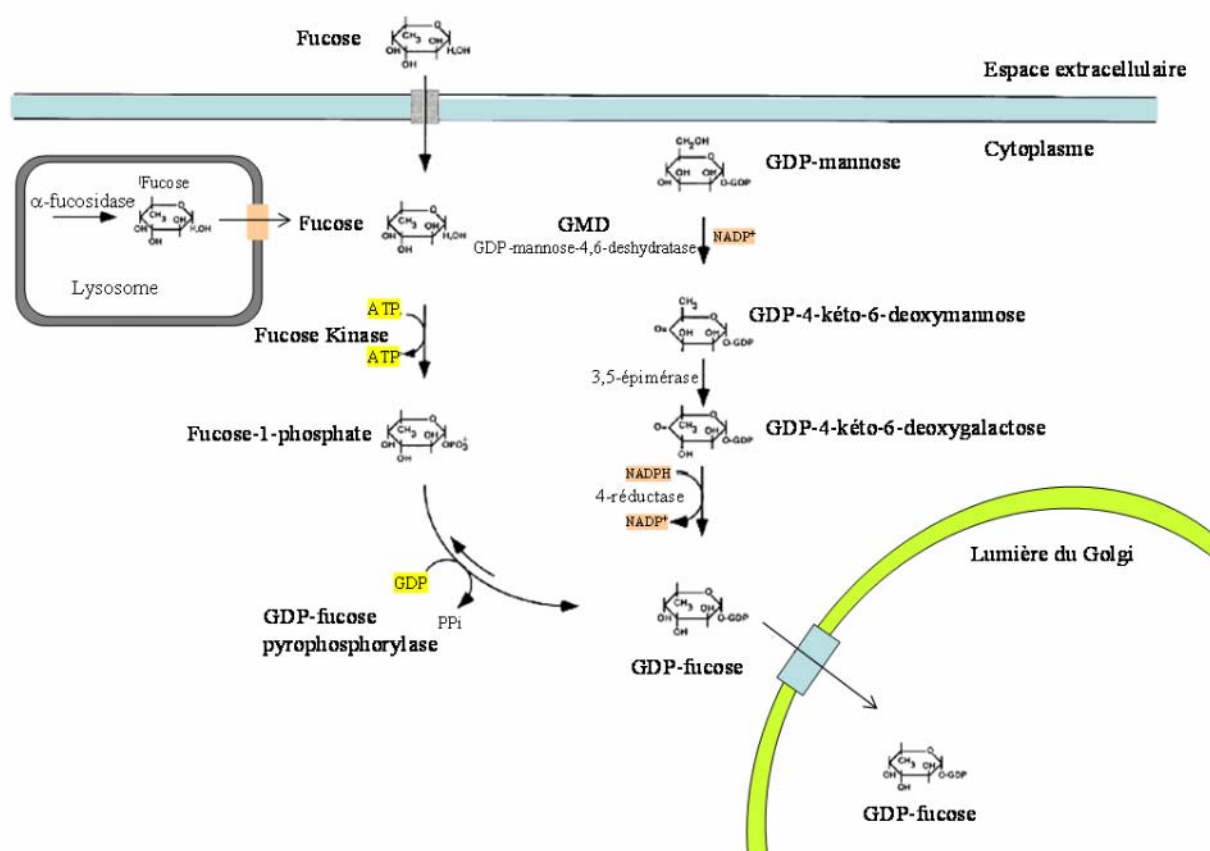


Figure 23 : Biosynthèse du GDP-fucose.

Chez les Mammifères, le GDP-fucose peut être synthétisé par deux voies. La voie *de novo*, à droite, est caractérisée par la conversion progressive du GDP-mannose en GDP-fucose. La deuxième voie dite de sauvetage permet de former du GDP-fucose directement à partir de fucose libre. Le GDP-fucose synthétisé est alors transporté depuis le cytoplasme vers la lumière de l'appareil de Golgi, via le transporteur de GDP-fucose golgien schématisé ici par un rectangle bleu (d'après Becker et Lowe, 2003).

Les enzymes Pofut1 et Pofut2 présentent entre elles peu d'homologie de séquences (à titre d'exemple, 23% d'identité chez l'Homme). Néanmoins, elles possèdent des caractéristiques communes telles qu'une séquence signal d'entrée dans le RE, la présence de sites potentiels de N-glycosylation et un motif DXD. Ce dernier se retrouve chez de nombreuses glycosyltransférases et serait impliqué dans la reconnaissance du groupement phosphate du nucléotide-sucres donneur (Gotting *et al.*, 2004) (Fig.24.).

Ces deux enzymes ont également des caractéristiques communes avec d'autres fucosyltransférases. En effet, Pofut1, Pofut2, les α 2- et les α 6-fucosyltransférases ont en commun trois motifs peptidiques conservés (I, II et III) (Martinez-Duncker *et al.*, 2003)

(Fig.24 et 25.) impliqués dans la catalyse enzymatique (Breton *et al.*, 1998; Chazalet *et al.*, 2001). Il est proposé que les gènes codant toutes ces fucosyltransférases soient le résultat d'événements de duplication à partir d'un ancêtre commun et qu'une évolution divergente soit à l'origine des trois familles d'enzymes reconnues aujourd'hui (Martinez-Duncker *et al.*, 2003).

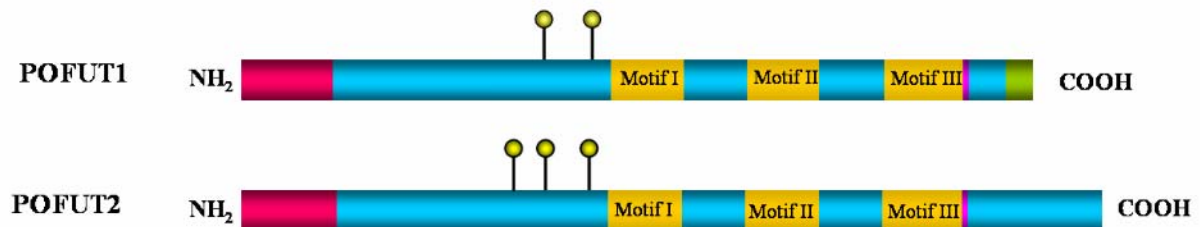


Figure 24 : Représentation schématique des enzymes POFUT1 et POFUT2 humaines. Le peptide signal est en rose, les motifs conservés entre les α 2-, α 6- et O-fucosyltransférases en orange, les sites potentiels de N-glycosylation en jaune, le DXD en violet et le KDEL-like de POFUT1 en vert.

				Motifs peptidiques conservés		
		Numéros				
Protéines	Espèces	d'accès	I	II	III	
α 2-fucosyltransférases	FUT1	<i>H. sapiens</i>	M35531	213	TFVGVHVRRGDYL<30>PFFVWVSSNGMEWCRK<20>WKDFALLTCQNHITMTI-GTFGFMAAYL	317
	"	<i>M. musculus</i>	U90553	215	TFVGVHVRRGDYL<30>PFFVWVSSNGMKWCLL<20>GKDFALLTCQNHITMTI-GTFGFMAAYL	319
	"	<i>S. scrofa</i>	L50534	213	TFVGVHVRRGDYL<30>PFFVWVSSNGMEWCRK<20>ARDFALLTCQNHITMTI-GTFGFMAAYL	317
	"	<i>B. taurus</i>	AF186465	213	TFVGVHVRRGDYL<30>PFFVWVSSNGMKWCRK<20>NKDFALLTCQNHITMTI-GTFGFMAAYL	317
	FUT2	<i>H. sapiens</i>	U17894	194	TFVGVHVRRGDYL<30>LFFVWVSSNGMAWCRK<20>AKDFALLTCQNHITMTI-GTFGFMAAYL	298
	"	<i>M. musculus</i>	AF214656	198	TFVGVHVRRGDYL<30>PFFVWVSSNGMAWCRK<20>AKDFALLTCQNHITMTI-GTFGFMAAYL	302
	"	<i>S. scrofa</i>	U70881	191	TYVGVHVRRGDYL<30>PFFVWVSSNGMAWCRK<20>AKDFALLTCQNHITMTI-GTFGFMAAYL	295
	"	<i>B. taurus</i>	X99620	195	TYVGVHVRRGDYL<30>PFFVWVSSNGMAWCRK<20>AKDFALLTCQNHITMTI-GTFGFMAAYL	299
	α 2FucT	<i>X. laevis</i>	E25301	207	TFTAIHVRRGDYL<30>PFFVWVSSNGMRWCRK<20>GKDFALLTCQNHITMTI-GTFGFVWSTL	311
	"	<i>C. elegans</i>	F 017784	209	FITCVHVRRGDYL<28>TATVFFGDDYEFMDS<21>LLYAKSNCDVVLITAAH-STFGWMMGYP	311
"	<i>C. elegans</i>	J 001660	204	FKTCVHIRRGDYL<25>GVYLFSDNDFGWVRS<23>LYFAOHCDAVLITAPS-STFGWMLGHL	306	
α 6-fucosyltransférases	FUT8	<i>H. sapiens</i>	D89289	358	PVIGVHVRRLDKV<32>RVYLATDDPSLL-KE<34>ILLDHLSCADPFVQTE-SSQVCRVAYE	477
	"	<i>M. musculus</i>	AB025198	358	PVIGVHVRRLDKV<32>RVYLATDDPTLL-KE<34>ILLDHLSCADPFVQTE-SSQVCRVAYE	477
	"	<i>S. scrofa</i>	D86723	358	PVIGVHVRRLDKV<32>RVYLATDDPAL-KE<34>ILLDHLSCADPFVQTE-SSQVCRVAYE	477
	"	<i>B. taurus</i>	AF247186	358	PVIGVHVRRLDKV<32>RVYLATDDPSLL-KE<34>ILLDHLSCADPFVQTE-SSQVCRVAYE	477
	"	<i>X. laevis</i>	AJ514872	361	PVIGVHVRRLDKV<32>RVYLATDDPTLL-KE<34>ILLDHLSCADPFVQTE-SSQVCRVAYE	480
	"	<i>D. melanog.</i>	AJ514872	361	PVIGVHVRRLDKV<32>RVYLATDDPTLL-KE<34>ILLDHLSCADPFVQTE-SSQVCRVAYE	480
	"	<i>C. elegans</i>	AJ512486	344	PVIGVHVRRLDKV<33>RIFLASDDPTV-PE<34>ILLDHLSCADPFVQTE-SSQVCRVAYE	464
	"	<i>C. intest. 2</i>	AJ515151	269	PVIGVHVRRLDKV<34>RIFLASDDPTV-PE<34>ILLDHLSCADPFVQTE-SSQVCRVAYE	389
	"	<i>C. intest. 5</i>	AJ515152	312	PVIGVHVRRLDKV<33>RIFLASDDPTV-PE<34>ILLDHLSCADPFVQTE-SSQVCRVAYE	431
	O-fucosyltransférases	POFUT1	<i>H. sapiens</i>	AF375884	233	PVYGIHLRNGSDW<59>SVYIATDSESYV-SE<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE
"		<i>M. musculus</i>	BC018194	287	PVYGIHLRNGSDW<59>SVYIATDSESYV-SE<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	370
"		<i>S. scrofa</i>	AJ567917	225	PVYGIHLRNGSDW<59>SVYIATDSESYL-PE<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	357
"		<i>G. gallus</i>	AJ535754	225	PVYGIHLRNGSDW<59>SVYIATDSEPYT-KE<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	357
"		<i>X. laevis</i>	AJ514425	225	PVYGIHLRNGSDW<59>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	357
"		<i>D. melanog.</i>	Q9V6X7	238	PVYGIHLRNGSDW<62>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	374
"		<i>A. gambiae</i>	EAA05278	229	PVYGIHLRNGSDW<62>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	365
"		<i>C. intestina</i>	AK112708	226	PVYGIHLRNGSDW<58>SLFIATDSESYE-NE<22>MDIALLIESDRFTGSCGSSVAFVSRK	360
"		<i>C. elegans</i>	CAD45602	365	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	496
POFUT2		<i>H. sapiens</i>	AJ575591	287	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	396
"		<i>M. musculus</i>	BC018194	287	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	396
"		<i>R. norvegic.</i>	XP_228073	287	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	396
"		<i>B. Taurus</i>	AJ575655	287	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	396
"		<i>G. gallus</i>	AJ575657	238	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	347
"		<i>C. elegans</i>	Q8WR51	283	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	391
"	<i>C. intestina</i>	AJ575656	283	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	392	
"	<i>D. melanog.</i>	Q9W589	264	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	376	
"	<i>A. gambiae</i>	EAA06818	224	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	437	
"	<i>P. yoelii</i>	EAA22665	278	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	387	
"	<i>P. falciparum</i>	Q8I360	272	PVYGIHLRNGSDW<57>SVYIATDSESYT-AP<19>QVDIYLLGQADHFIGNCVSSFTAFVKRE	381	

Figure 25 : Alignement des motifs conservés d' α 2-, α 6- et O-fucosyltransférases. Les acides aminés conservés à 70% dans au moins deux familles sont représentés sur un fond noir, ceux conservés à 70% dans une seule famille sont représentés sur un fond gris. Les acides aminés [VILM], [YFW], [RKH] et [EDQN] sont considérés comme étant équivalents du point de vue de leurs propriétés physico-chimiques (extrait de Martinez-Duncker *et al.*, 2003).

2.4 Existe t'il une troisième O-fucosyltransférase ?

Du O-fucose a été trouvé sur l'inhibiteur de protéinase, PMP-C (Pars intercerebralis Major Peptide C), chez le criquet migrateur *Locusta migratoria*. La PMP-C possède 6 cystéines conservées comme les EGF et les TSR; cependant, elles sont impliquées dans la formation de 3 ponts disulfures différents (C¹-C⁴/C²-C⁶/C³-C⁵). De plus, la séquence cible de la O-fucosylation est distincte des séquences consensus précédemment décrites (Nakakura *et al.*, 1992). Très peu de données sont disponibles sur la O-fucosylation de cette protéine, mis à part le fait qu'elle serait impliquée dans sa stabilisation (Mer *et al.*, 1996). Il n'a pas été déterminé si Pofut1, Pofut2 ou une nouvelle O-fucosyltransférase était responsable de la O-fucosylation de cette protéine. Si une telle enzyme existe, elle doit être suffisamment divergente des deux premières pour expliquer qu'on ne puisse l'identifier *in silico* sur la base d'homologie de séquences.

2.5 O-fucosylation et développement embryonnaire

La O-fucosylation des domaines EGF, la plus étudiée aujourd'hui, est une modification post-traductionnelle essentielle au fonctionnement correct de plusieurs protéines, dont Notch et ses ligands (Moloney, Panin *et al.*, 2000; Moloney, Shair *et al.*, 2000; Panin *et al.*, 2002), eux mêmes impliqués dans de nombreux processus développementaux telle que la régulation du processus myogénique qui nous intéresse plus particulièrement.

En 2003, des observations phénotypiques et moléculaires sur des souris KO pour le gène *Pofut1*, créées par Shi et Stanley, ont confirmé le rôle prédominant de la O-fucosylation dans le développement embryonnaire. En effet, les souris KO meurent au stade embryonnaire E9,5 suite à de nombreux dysfonctionnements dans divers processus développementaux.

Les souris KO pour le gène *Pofut1* présentent, en effet, un important retard développemental dans les processus de vasculogénèse et de cardiogénèse. A titre d'exemple, au stade E8,25, leur coeur est beaucoup plus petit que celui des souris témoins et leur croissance s'arrête au stade E8,75. Ces souris sont également affectées au niveau de la neurogénèse. Leur tube neural n'est pas rectiligne et présente de nombreuses circonvolutions.

La malformation qui nous intéresse le plus, étant donné le cadre de nos travaux, est un important défaut de formation des somites. Comme nous le verrons dans le chapitre 3, les somites sont des structures embryonnaires à partir desquelles dérivent, entre autres, les cellules des muscles squelettiques. Au stade E8,5, les souris *Pofut1* *-/-* sont de taille équivalente aux souris sauvages. Néanmoins, les somites postérieurs adjacents au mésoderme pré-somitique sont fusionnés. D'autres somites situés dans la partie antérieure des embryons sont également fusionnés et sont de taille et de forme variables. Au stade E8,75, des expériences d'hybridation *in situ* montrent que des transcrits codant le récepteur Notch1 sont

exprimés comme chez l'embryon de souris sauvage, bien que de façon plus diffuse (Fig.26a.), alors que le gène *Hes5* (un des gènes activés par les cascades de signalisation de Notch) est exprimé de façon moindre par rapport à un embryon sauvage (Fig.26b.). Ce résultat montre une non fonctionnalité du récepteur Notch1 lié à l'absence de *O*-fucose sur ses domaines EGF. Par contre, au stade E9, l'expression des transcrits *Notch1* a diminué et est plus faible que chez une souris sauvage (Fig.26c.), confirmant que le récepteur Notch1 autorégule sa propre expression durant la somitogenèse (Barrantes *et al.*, 1999). De même, l'expression des transcrits codant les ligands de Notch, Jagged1 (*Jag1*) et Delta3 (*Dll3*), ainsi que ceux codant Lunatic Fringe (*Lfng*), est plus faible au niveau des somites chez les souris KO (Fig.26d, e et f.) alors que dans d'autres tissus comme le cerveau ou le tube neural, *Dll3* ou *Lfng* sont respectivement surexprimés (Fig.26 e et f.). De tels résultats révèlent l'existence de mécanismes complexes de rétrocontrôles, négatifs et positifs, s'exerçant entre les différents intervenants des voies de signalisation par Notch.

Au stade E9,5, les embryons murins KO pour le gène *Pofut1* meurent. Ils présentent un important retard développemental avec, entre autres, un nombre de somites largement inférieur à la normale. De plus, ces embryons présentent une expression moindre des transcrits codant la Myogénine (Fig.26g.) qui est un facteur de transcription impliqué dans la détermination et la différenciation des cellules musculaires. Ainsi, l'analyse de son niveau d'expression est souvent utilisée comme marqueur du myotome, portion des somites contenant les cellules souches musculaires.

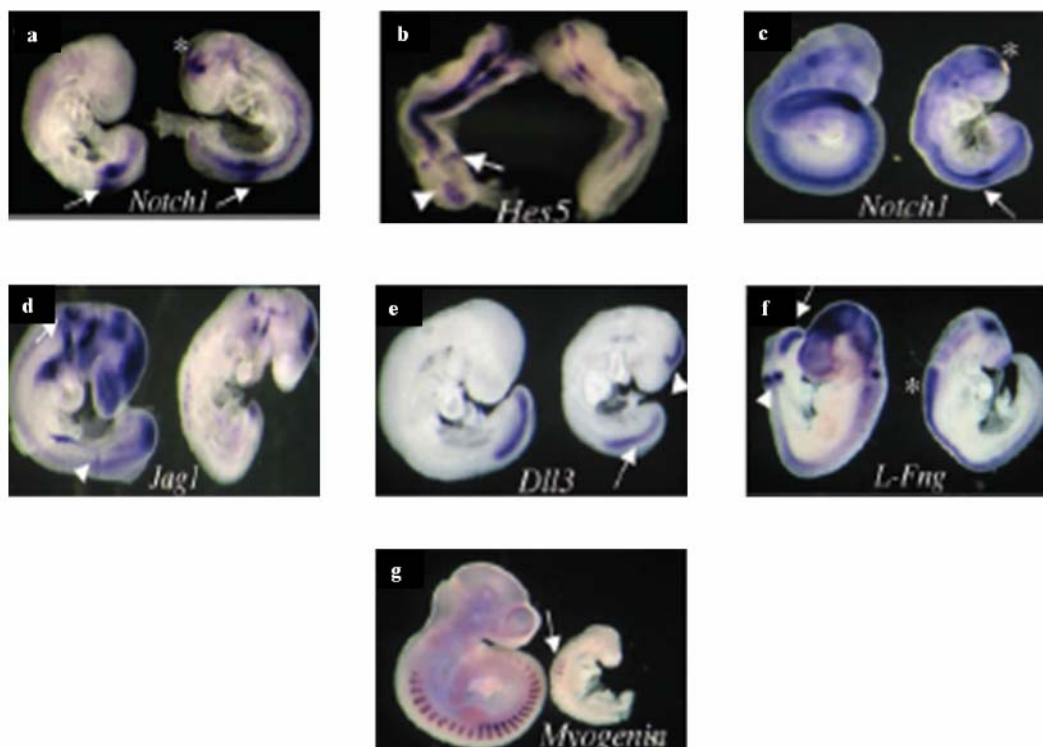


Figure 26 : Expression des gènes impliqués dans les voies de signalisation par Notch chez les embryons de souris sauvages (gauche) et mutants *Pofut1*^{-/-} (droite).

a : *Notch1* au stade E8,75; l'expression est la même au niveau des somites adjacents au mésoderme pré-somitique bien que plus diffuse chez l'embryon KO. b : *Hes5* au stade E8,5; chez l'embryon sauvage, le gène est exprimé dans le système nerveux et les somites, mais son expression est très diminuée chez l'embryon KO. c : *Notch 1* au stade E9; l'expression est plus forte au niveau du mésoderme pré-somitique chez l'embryon sauvage alors qu'elle est plus importante dans le cerveau de l'embryon KO (indiqué par une étoile). d, e et f : *Jag1*, *Dll3* et *Lfng* sont régulés négativement au niveau des somites postérieurs de l'embryon KO. e : La surexpression de *Dll3* dans le cerveau de l'embryon KO est indiquée par une tête de flèche. f : La surexpression de *Lfng* dans le tube neural de l'embryon KO est indiquée par une étoile. g : *Myogénine* au stade E9,5; des transcrits *Myogénine* sont très faiblement détectés au niveau des quelques somites que possède l'embryon KO. Le retard développemental de ce dernier est flagrant. En effet, sa taille n'atteint pas le tiers de celle d'un embryon murin sauvage (extrait de Shi et Stanley, 2003).

L'enzyme O-fucosyltransférase 1 est donc absolument nécessaire au développement et à la survie d'un organisme, en particulier, par son action sur les membres de la famille Notch. Les phénotypes des souris KO pour le gène *Pofut1* sont, en effet, les mêmes que ceux générés par des KO de gènes codant des protéines intervenant en aval des voies de signalisation des quatre récepteurs Notch présents chez les Mammifères. L'enzyme Pofut1 participe donc au développement et à la segmentation correcte des somites, nécessaire à leur devenir myogénique.

Pour l'instant, aucune protéine impliquée dans le processus myogénique n'a été décrite comme possédant des domaines TSR *O*-fucosylés ou même *O*-fucosylables. Cependant, il est intéressant de noter que les expériences d'ARN interférence et de surexpression du gène *Pofut2* chez *C. elegans* (Menzel *et al.*, 2004) ont provoqué de nombreuses malformations avec notamment, une mauvaise compartimentation et localisation des précurseurs musculaires (body wall muscle cells), équivalents des cellules somitiques des Mammifères. La *O*-fucosylation de type 2 pourrait donc jouer un rôle au cours du développement musculaire.

Pourquoi glycosylation et myogenèse ?

Un des aspects de mon travail de thèse a consisté à rechercher le lien unissant la *O*-fucosylation et la différenciation de la cellule musculaire. La myogenèse est un processus extrêmement complexe, nécessitant de nombreuses interactions cellulaires. La *O*-fucosylation n'est évidemment pas la seule modification post-traductionnelle y jouant potentiellement un rôle important.

Ainsi, l'unité de recherche dans laquelle j'ai réalisé mes travaux de thèse, s'intéresse aux rôles joués par les glycoconjugués dans les étapes de prolifération et de différenciation de la cellule musculaire. Plusieurs données bibliographiques récentes soulignent l'importance fonctionnelle des glycoprotéines et des protéoglycane dans les processus myogéniques et la physiologie du muscle squelettique. Certaines pathologies neuromusculaires sont, en effet, dues à des mutations dans des gènes de glycosyltransférases.

Cette partie de l'exposé bibliographique est dédiée à la présentation de quelques exemples soulignant l'intérêt d'établir une relation étroite entre la glycosylation et le développement ou la physiologie du muscle.

> La N-glycosylation est importante pour la fusion des cellules musculaires

Bien que les données sur l'importance des N-glycanes lors de la myogenèse soient très peu nombreuses, nous savons que durant les processus de fusion cellulaire conduisant à terme à la formation des myotubes et des fibres musculaires (cf. Chapitre suivant), la composition en N-glycoprotéines de la surface cellulaire est modifiée.

En 1986, Holland et Herscovics ont utilisé sur des myoblastes L6 en culture un inhibiteur de glucosidases (N-méthyl-1-deoxynojirimycine ou MDJN) ou un inhibiteur de mannosidases (1-deoxymannojirimycine ou ManDJN), ceux-ci inhibent la formation des N-glycanes complexes en empêchant la formation du précurseur : $\text{Man}_8\text{GlcNAc}_2$. Les auteurs ont montré que l'utilisation du MDJN bloque la fusion des myoblastes alors que l'utilisation du ManDJN a peu d'effet. Ils en ont déduit que les N-glycanes complexes ne sont pas nécessaire à la fusion myoblastique alors que les N-glycanes oligomannosidiques doivent avoir un rôle. L'absence de fusion des cellules traitées avec le MDJN est sans doute dû à la présence sur des glycoprotéines de surface de glycanes $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$, les glucoses qui normalement ne devraient pas être présents empêchent sans doute la reconnaissance entre des protéines de surface dont les interactions sont nécessaires à la fusion. En 1984, Cates et ses collaborateurs ont isolé un mutant spontané de cette lignée incapable de fusionner pour former des myotubes. Il s'est avéré que ce mutant était déficient dans le transfert des résidus mannoses sur le dolicholpyrophosphate et ne pouvait donc pas assurer la biosynthèse normale des N-glycanes.

Une étude plus récente montre que des souris KO pour le gène *Mgat1*, codant la GlcNAc transférase I (GlcNAc-TI), impliquée dans la première étape de la formation des N-glycanes complexes et hybrides, meurent *in utero* au bout du dixième jour (E10) suite à des malformations du tube neural et à un mauvais établissement des axes embryonnaires (Metzler *et al.*, 1994). Cette souris est beaucoup plus petite que la souris sauvage et présente un retard dans le développement des somites (20 contre 26 somites chez la souris sauvage à E9,5).

> Les protéoglycanes interviennent dans la myogenèse à plusieurs niveaux

Les cellules des muscles squelettiques sont recouvertes d'une importante matrice extracellulaire (MEC). D'une architecture complexe, la MEC est formée de glycoprotéines, de collagène et de protéoglycanes. Outre son rôle dans la plasticité du muscle

adulte, l'expression spatio-temporelle de ses constituants crée un micro-environnement autour des cellules musculaires, favorisant ou inhibant ses interactions avec des ligands potentiels et/ou avec des cellules voisines.

Les protéoglycanes interagissent avec les fibres de collagène et ainsi contribuent à l'organisation de la MEC. Ils interviennent dans la modulation des signaux extérieurs (ligands) destinés à la cellule. Au sein de la MEC, les protéoglycanes jouent des rôles importants dans le processus myogénique en modulant, par exemple, la prolifération cellulaire, la réponse aux facteurs de croissance et la fibrillogenèse (Velleman, 1999). Ainsi, la composition en protéoglycanes de la MEC change au cours du développement musculaire. On observe en effet, une transition entre une MEC riche en chondroïtines sulfates au début du développement et une MEC riche en dermatanes et héparanes sulfates lors de la fusion des myoblastes. Ce basculement dans la composition de la MEC a été observé chez le poulet (Fernandez *et al.*, 1991), la souris (Young *et al.*, 1990) et la dinde (Velleman *et al.*, 1999). Il semblerait donc que les différents types de protéoglycanes présents à la surface des cellules musculaires aient des fonctions spécifiques durant la myogenèse.

La capacité du tissu musculaire à se régénérer en cas de blessure est une propriété liée à la présence, entre la lame basale et la fibre musculaire, de cellules mononucléées, les cellules satellites. A la naissance, les cellules satellites comptent pour environ 32% des cellules musculaires contre 5% chez l'adulte. Ces cellules sont quiescentes, mais la réception de signaux extérieurs provenant du tissu lésé va leur permettre de proliférer à nouveau et de se différencier en myotubes pour former de nouvelles fibres musculaires. Les mécanismes gouvernant la différenciation de ces cellules ressemblent à ceux intervenant au cours du développement embryonnaire. Ils sont modulés à la fois par des interactions cellule-cellule, cellule-MEC et MEC-facteurs de croissance. Zimowska et ses collaborateurs ont montré en 2001, qu'un traitement par des agents capables de mimer des protéoglycanes à héparane sulfate accélère le processus de régénération d'un muscle de rat, *in vivo*, après blessure et dénervation. De plus, *in vitro*, ce traitement augmente la prolifération et la différenciation de cellules satellites musculaires de rat et de cellules myoblastiques murines de la lignée C2.7 (Papy-Garcia *et al.*, 2002). Les cellules myoblastiques murines de la lignée C2C12 cultivées en présence d'inhibiteur spécifique de la synthèse de certains des protéoglycanes à héparane et chondroïtine sulfate (héparitinase, chondroïtinase) ne peuvent plus se différencier en myotubes (Villena et Brandan, 2004). A l'inverse, un traitement de ces cellules par des protéoglycanes à héparane sulfate ou des structures capables de les mimer (RGTA,

ReGeneraTing Agents) accélère leur prolifération et leur différenciation (Zimowska *et al.*, 2001; Papy-Garcia *et al.*, 2002).

> La *O-N*-acétylglucosaminylation régule la physiologie du muscle

Comme nous l'avons vu dans le chapitre 1 de cet exposé bibliographique, la *O-N*-acétylglucosaminylation est largement répandue au sein des cellules eucaryotes soulignant un rôle physiologique fondamental inhérent à cette modification post-traductionnelle.

La *O-N*-acétylglucosaminylation est une modification réversible catalysée par l'enzyme OGT. Elle aurait un rôle régulateur similaire à celui du système de phosphorylation/déphosphorylation. De plus, ces deux modifications sont souvent réciproques sur un même acide aminé (S/T) d'une protéine. Cette relation étroite entre présence de *O*-GlcNAc ou de phosphate est nommée le Yin-Yang (Lefevre *et al.*, 2003). Au sein d'une cellule, le taux d'UDP-GlcNAc, substrat donneur de l'enzyme OGT, est régulé par le taux de glucose (Hanover *et al.*, 1999; Wells *et al.*, 2003). Les phénomènes de contraction musculaire sont également dépendants du métabolisme du glucose et sont, entre autres, régulés par des processus de phosphorylation/déphosphorylation. Ces observations ont conduit Cieniewski-Bernard et ses collaborateurs en 2004 à s'intéresser à la *O-N*-acétylglucosaminylation dans le muscle squelettique de rat. Les *O*-GlcNAc seraient impliquées dans la physiologie du muscle en contrôlant le métabolisme du glucose par la modulation de l'activité des enzymes de la glycolyse. Le passage d'un état oxydatif (lent) à un état glycolytique (rapide) d'un muscle s'accompagne, par exemple, d'une diminution du taux de *O*-GlcNAc d'environ 20,4%. La *O-N*-acétylglucosaminylation, intervenant sur des protéines spécifiques du muscle (α B-crystalline, chaînes de myosines), serait aussi impliquée dans le maintien de leur homéostasie et serait donc un élément régulateur de l'atrophie musculaire (Cieniewski-Bernard *et al.*, 2005).

> La *O*-mannosylation est impliquée dans la survenance de dystrophies musculaires

Assez récemment, les modifications post-traductionnelles des protéines sont devenues les cibles de nouvelles recherches sur les dystrophies musculaires. Plusieurs pathologies de ce type sont connues chez l'Homme : le WWS (Walker Warburg Syndrome), le MEB (Muscle-Eye-Brain disease), le FCMD (Fukuyama Congenital Muscular Dystrophy), les CMD1C et

1D (Congenital Muscular Dystrophy type 1C/1D), le LGMD2I (Limb Girdle Muscular Dystrophy type 2I) (Grewal et Hewitt, 2003) et chez la souris, le myd (Myodystrophy) (Grewal *et al.*, 2001). Toutes ces maladies possèdent des symptômes communs (dystrophie musculaire, malformation des yeux et du cerveau, parfois retard mental) et il est quelque fois difficile de les diagnostiquer avec précision. Cependant, elles ont comme point commun la perte de la *O*-mannosylation de la sous unité alpha du dystroglycane (α -dystroglycane), d'où leur nouveau nom de dystroglycanopathies.

Le dystroglycane est composé de deux sous unités, une α (*O*-mannosylée) et une β (*N*-glycosylée). Il est le composant majeur du complexe DGC (Dystrophin-associated Glycoprotein Complex) qui assure la liaison entre l'actine du cytosquelette et les composants de la matrice extracellulaire (MEC) (Fig .27.)

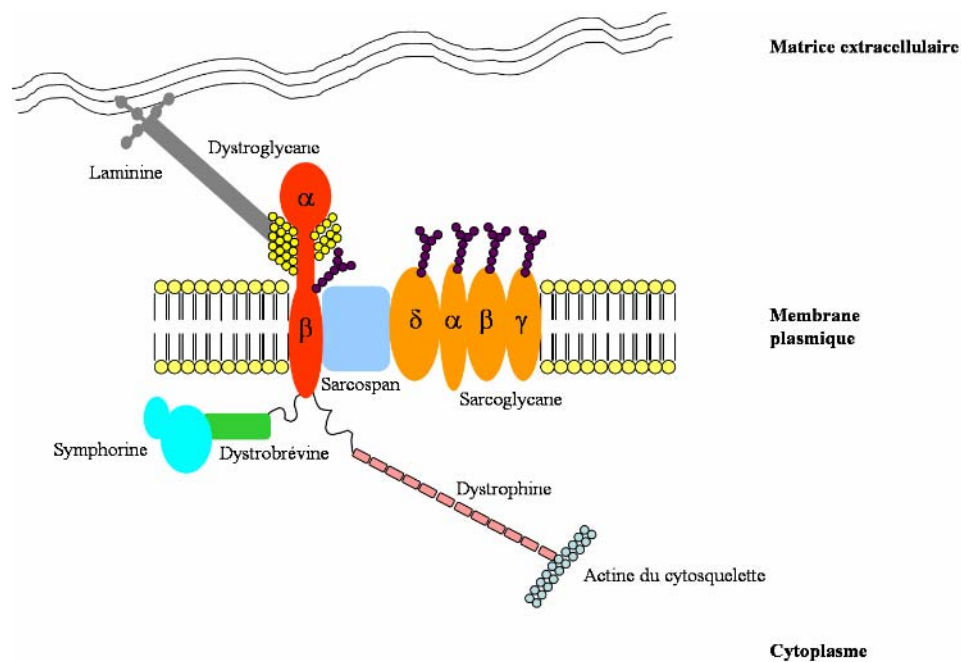


Figure 27 : Représentation schématique du complexe DGC.

Dans le muscle squelettique, le DGC est composé de Symphorine, de Dystrobrevine, de Dystrophine qui relie le complexe membranaire à l'actine du cytosquelette, des deux sous-unités du Dystroglycane, des quatre sous-unités du Sarcoglycane et du Sarcospan. Les *N*-glycanes sont indiqués en violet et les *O*-mannosylglycanes en jaune. Plus de la moitié de la masse de l' α -dystroglycane est due à la présence de *O*-glycanes. Des mutations dans les gènes codant ces protéines résultent en une dystrophie musculaire. La laminine est un hétérodimère (sous unités α , β et γ non représentées) qui relie le DGC à la MEC par l'intermédiaire des *O*-mannosylglycanes. Extrait de Grewal et Hewitt, 2003.

Plusieurs équipes de recherche ont isolé et caractérisé les *O*-mannosylglycanes présents sur l' α -dystroglycane (Chiba *et al.*, 1997; Sasaki *et al.*, 1998; Smalheiser *et al.*, 1998) (cf. Chapitre 1).

Il a été démontré que des mutations dans les gènes *POMT1* et *POMGnT1* étaient responsables de certaines des dystroglycanopathies. En effet, une mutation dans le gène *POMT1* est détectée dans 20% des patients atteints du WWS (De Bernabe *et al.*, 2003). Un deuxième gène codant potentiellement une autre *O*-mannosyltransférase, *POMT2*, a été identifié (Willer *et al.*, 2002). Peu de données sur ce gène sont actuellement disponibles, néanmoins sa mutation pourrait être également responsable de la survenance du syndrome de Walker-Warburg (Van Reeuwijk *et al.*, 2005), bien que celui-ci semble faiblement exprimé dans le muscle squelettique (Willer *et al.*, 2002). De même, des mutations dans le gène *POMGnT1* ont été décrites chez des patients atteints du MEB (Yoshida *et al.*, 2001; Zhang, Betel *et al.*, 2002).

La conséquence d'une telle atteinte de l' α -dystroglycane est une fragilisation puis probablement une altération de la membrane plasmique au fil du temps par les contractions répétées, d'où une intolérance à l'effort (cf. Association française contre la myopathie www.afm-france.org), ce qui entraîne un décès à un âge prématuré.

Le *O*-mannosylglycane de structure : NeuAc α 2,3Gal β 1,4GlcNAc β 1,2Man α 1-*O*-S/T ressemble beaucoup au *O*-fucosylglycane : NeuAc α 2/3,6Gal β 1,4GlcNAc β 1,3Fuc α 1-S/T, trouvé sur les domaines EGF de certaines protéines. Les enzymes responsables de leur allongement en tri- et tétrasaccharide pourraient être les mêmes. Ainsi, les données actuellement disponibles sur l'implication des *O*-mannosylglycanes dans certaines pathologies dystrophiques renforcent l'hypothèse du rôle possible des *O*-fucosylglycanes dans le processus de myogenèse. De même, enrichir les connaissances sur la formation et le rôle des *O*-fucosylglycanes pourrait permettre, à terme, de mieux comprendre l'implication des *O*-mannosylglycanes dans le développement et la physiologie du muscle squelettique.

La glycosylation, et plus particulièrement la O-fucosylation, intervient sur des protéines aux rôles biologiques fondamentaux, notamment durant l'embryogenèse précoce. Le greffage de O-fucose sur des protéines majeures (Notch et ses ligands, par exemple) serait déterminant lors de la mise en place des axes embryonnaires et la détermination des cellules souches dans diverses voies de différenciation. Au cours de ma thèse, je me suis intéressée à la voie de différenciation de la cellule musculaire inscrite dans le mécanisme de Myogenèse. Ce processus hautement régulé fait l'objet du troisième chapitre de ce manuscrit.

3 La myogenèse du muscle squelettique

3.1	ORIGINE EMBRYONNAIRE DU MUSCLE SQUELETTIQUE.....	93
3.2	DETERMINATION MYOGENIQUE.....	97
3.2.1	<i>Détermination et migration des prémyoblastes</i>	97
3.2.2	<i>Différenciation des précurseurs musculaires</i>	99
3.3	EFFECTEURS DE LA MYOGENESE	107
3.3.1	<i>Les inducteurs</i>	107
3.3.2	<i>Les inhibiteurs</i>	109
3.4	LA MYOGENESE <i>IN VITRO</i>	115

Trois principaux types de muscles sont présents chez les Vertébrés et assurent des fonctions relatives à la contraction cellulaire. Traditionnellement, ces trois types musculaires, cardiaque, lisse et squelettique, sont distingués par leurs attributs structuraux et fonctionnels. Dans cette partie, nous détaillerons seulement la formation du muscle squelettique. Le développement du muscle squelettique lors de l'embryogenèse est un processus complexe qui consiste en une séquence d'évènements cellulaires incluant : la détermination des cellules souches du mésoderme somitique ou paraxial en cellules de la lignée musculaire squelettique, la prolifération des myoblastes précurseurs, leur sortie du cycle cellulaire et leur fusion en myotubes polynucléés exprimant des protéines spécifiques du muscle squelettique différencié, notamment celles du système contractile.

3.1 Origine embryonnaire du muscle squelettique

Chez les Vertébrés, hormis quelques muscles de la tête qui dérivent du mésoderme de la tête et du mésoderme pré-chordal, les cellules précurseurs myogéniques des muscles squelettiques dérivent de cellules somitiques multipotentes (Eloy-Trinquet et Nicolas, 2002). Les somites sont des structures embryonnaires sphériques transitoires situés de part et d'autre du tube neural et de la notochorde, issues de la segmentation du mésoderme paraxial suivant un axe antéro-postérieur (Fig.28.).

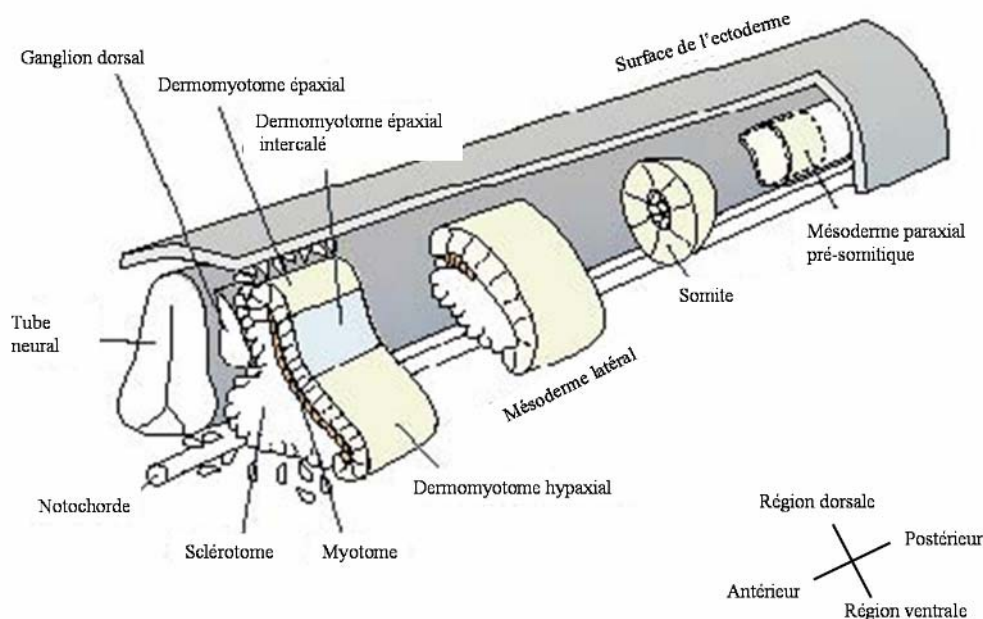


Figure 28 : Représentation schématique de la somitogenèse chez les Vertébrés. D'après Buckingham *et al.*, 2003.

Peu de temps après leur formation, les somites vont se compartimenter en deux parties, ventrale et dorsale. La partie ventrale donne le sclérotome, à l'origine des vertèbres et des côtes (Kalcheim *et al.*, 1999). La partie dorsale, aussi appelée dermomyotome, abrite les

précurseurs de la lignée dermique du dos (dermatome) et de la lignée myogénique (myotome). Cette partie des somites subdivisée en deux régions, médiane et latérale, est composée des précurseurs de deux lignées myogéniques (Ordahl et Le Douarin, 1992). Dans la partie médiane, les cellules restent dans le territoire somitique puis se différencient pour donner la musculature paraxiale à l'origine des muscles intercostaux et paraspinaux associés à la colonne vertébrale. Dans la moitié latérale, les précurseurs sont destinés à quitter le territoire des somites pour migrer vers leur localisation finale où ils formeront la musculature des membres et de l'abdomen (Christ et Ordahl, 1995; Brand-Saberi et Christ, 1999) (Fig.29).

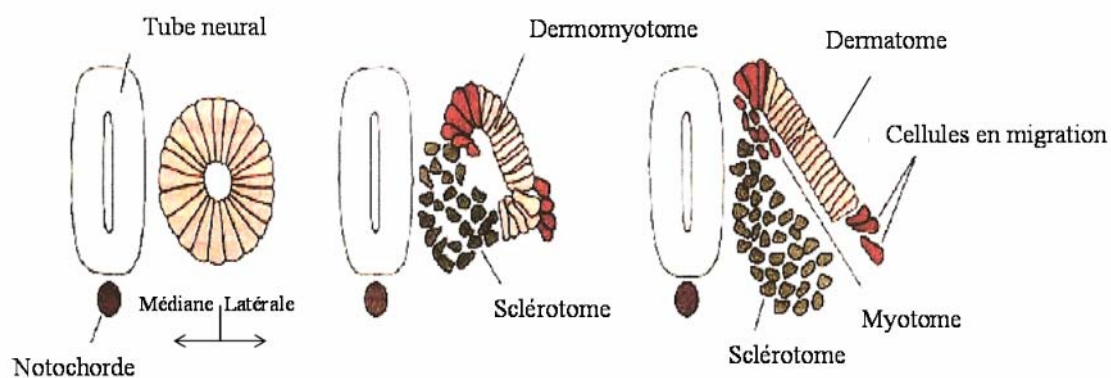


Figure 29 : Représentation schématique de la somitogenèse chez les Vertébrés.

Les cellules mésodermiques se condensent pour former les somites qui donnent deux types de tissus : le sclérotome et le dermomyotome. Au final, à partir des somites, quatre types cellulaires vont se différencier : les cellules du dermatome qui donneront les cellules dermiques du dos, les cellules du myotome qui après migration donneront la musculature des membres et de l'abdomen, les cellules du myotome qui donneront la musculature axiale, et les cellules du sclérotome qui donneront les côtes et les vertèbres.

Les muscles de la région des membres sont donc formés à partir de cellules précurseurs migratrices (Buckingham *et al.*, 2003) issues de la partie ventro-latérale du dermomyotome. Ces cellules somitiques dites « précurseurs » subissent un processus de détermination qui les fait évoluer en cellules fusiformes, dénommées myoblastes présomptifs ou prémyoblastes.

Chez les Vertébrés, après s'être détachés et avoir migré des somites jusqu'aux sites de formation des muscles, les prémyoblastes prolifèrent, expriment différents gènes régulateurs spécifiques du muscle et deviennent des myoblastes (Kenny-Mobbs et Thorogood, 1987, Sabourin et Rudnicki, 2000). Arrivés au sein des masses prémusculaires, l'activité mitotique cesse. Les myoblastes quittent le cycle cellulaire de manière irréversible et fusionnent pour

former des myotubes embryonnaires primaires. Ce mécanisme résulte de plusieurs évènements successifs : alignement des myoblastes selon un axe longitudinal, reconnaissance et adhésion intercellulaires, union des bicouches lipidiques entre myoblastes adjacents afin d'établir une continuité cytoplasmique entre les cellules (Wakelam, 1985).

Les myotubes se différencient et accumulent les protéines contractiles, actine et myosine, qui sont arrangées en faisceaux, les myofibrilles. L'arrangement de ces protéines contractiles en filaments fins d'actine et en filaments épais de myosine, sera à l'origine des unités contractiles répétées visibles le long des myofibrilles, les sarcomères. Initialement, ces myofibrilles sont situées à la périphérie des myotubes nouvellement formés car les différents noyaux sont localisés en position centrale. Ces derniers sont par la suite repoussés vers la périphérie du myotube par les myofilaments qui s'organisent au centre selon l'axe longitudinal de la cellule. Les myotubes accumulent également des protéines régulatrices et contractiles comme la tropomyosine et la troponine, ainsi que des enzymes cytoplasmiques nécessaires à la production d'énergie comme la myokinase et la phosphocréatine kinase.

Certaines cellules restent à l'état mononucléé entre les membranes basale et plasmique des fibres musculaires squelettiques et représentent un réservoir de cellules précurseurs myogéniques indifférenciées à l'état quiescent. Ce sont les cellules satellites. En réponse à un stimulus, ces cellules s'activent, prolifèrent et expriment les marqueurs myogéniques. Elles participent ainsi à la réparation et la régénération du muscle adulte.

La figure 30 résume les différentes étapes se produisant depuis la détermination jusqu'à la différenciation myogénique.

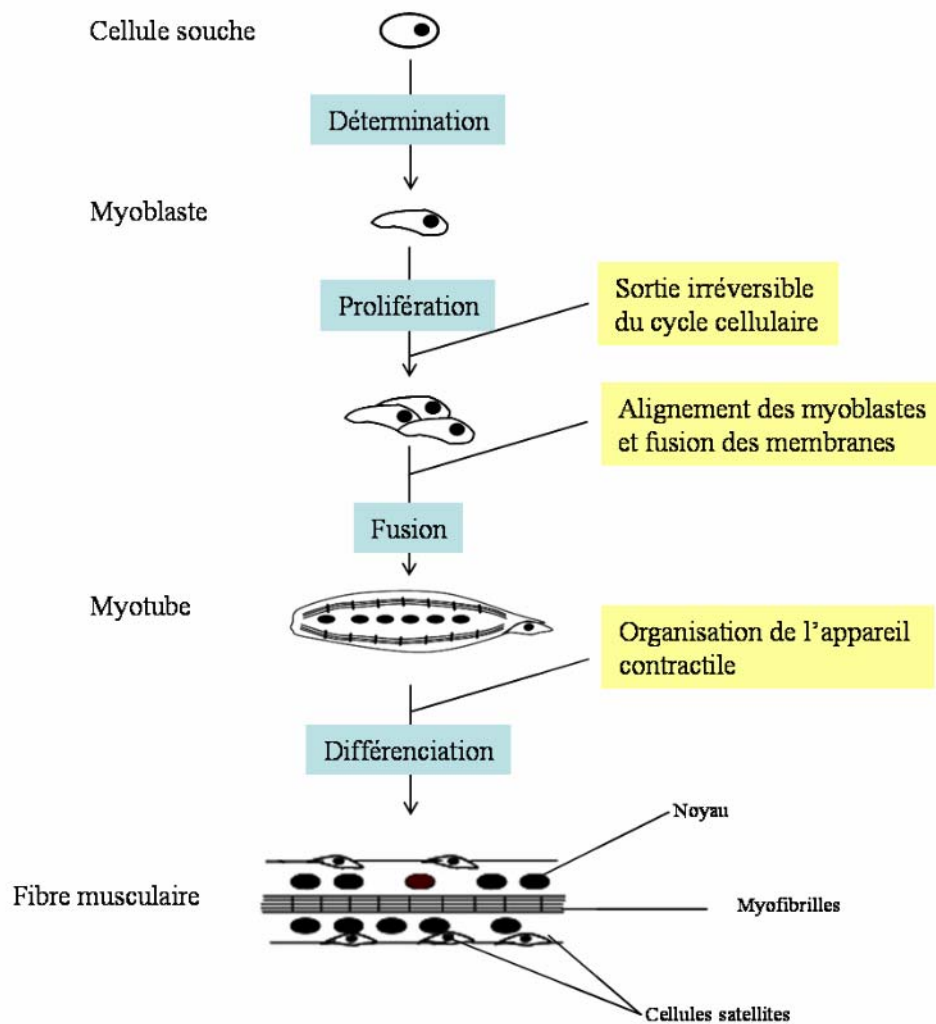


Figure 30 : Schéma illustrant la séquence des évènements qui aboutissent à la formation des fibres musculaires striées.

3.2 Détermination myogénique

De nombreux facteurs myogéniques interviennent dans la régulation de la myogenèse, du détachement des précurseurs à la différenciation en myotubes. Ils participent ainsi au contrôle du développement et à la mise en place du tissu musculaire squelettique.

3.2.1 Détermination et migration des prémyoblastes

Avant de se différencier, les cellules précurseurs sont donc localisées dans le dermomyotome (Venters *et al.*, 1999; Venters et Ordahl, 2002). Elles reçoivent des signaux dorsaux (détaillés plus loin), inducteurs du programme myogénique, c'est à dire entraînant l'expression des gènes de détermination myogénique (Fig.31.).

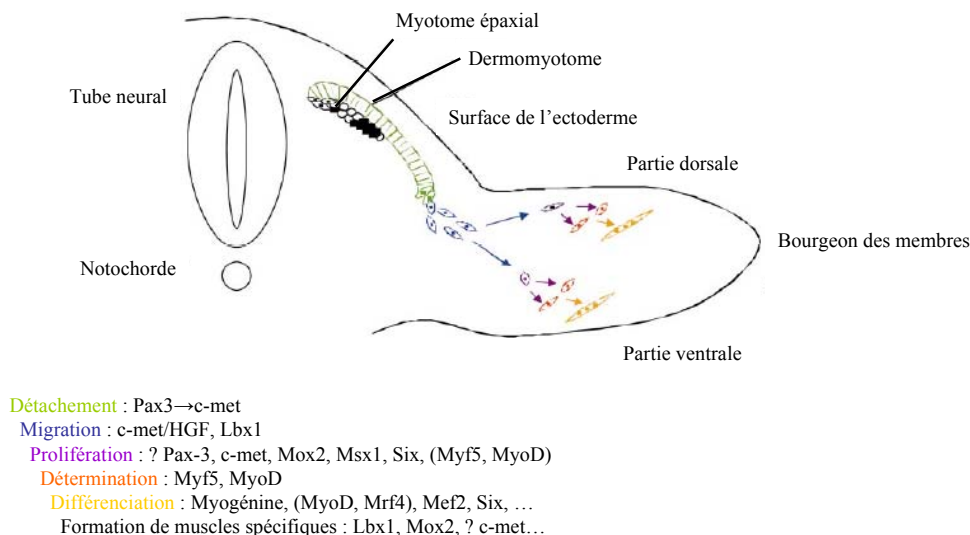


Figure 31 : Représentation schématique de la formation du muscle squelettique au niveau des membres chez les Vertébrés.

Les cellules précurseurs se détachent du dermomyotome, migrent vers les bourgeons des membres où elles prolifèrent et expriment différents facteurs myogéniques, avant de se différencier en cellules musculaires squelettiques. Chaque étape et les gènes qui y sont potentiellement associés sont représentés par différentes couleurs (d'après Buckingham *et al.*, 2003).

Les cellules du dermomyotome et les cellules migratrices expriment le gène *Pax3* appartenant à la famille multigénique *Pax* et codant un facteur de transcription à homéodomaine et un domaine paired assurant la fixation à l'ADN. Il joue un rôle fondamental dans la mise en place et le devenir du dermomyotome hypaxial, ainsi que dans le détachement et la migration des cellules progénitrices du muscle. Ainsi, des souris KO pour le gène *Pax3* sont dépourvues de muscles au niveau des membres (Bober *et al.*, 1994).

D'autres protéines sont impliquées dans la survie, le détachement et la migration des cellules musculaires précurseurs jusqu'aux sites de formation des muscles du corps et des membres (Fig.31.). Des expériences d'inactivation de gènes chez la souris ont permis d'identifier précisément leur rôle au cours de la somitogenèse (Tableau 2).

Tableau 2 : Protéines impliquées dans la détermination et la migration des précurseurs musculaires
D'après Buckingham *et al.*, 2003.

Gènes	Phénotypes des mutants/Fonctions potentielles
<i>Pax3</i> -/-	Désintégration du dermomyotome hypaxial, perte des cellules musculaires précurseurs, pas de détermination, pas de migration, pas de muscles des membres.
<i>Lbx1</i> -/-	Détachement des cellules musculaires précurseurs mais migration incorrecte, absence complète de certains muscles squelettiques.
<i>Mox2</i> -/-	Détachement des cellules musculaires précurseurs mais migration incorrecte, masse musculaire faible et présence de nombreuses fibres myopathiques.
<i>c-met</i> -/-; <i>HGF</i> -/-	Dermomyotome hypaxial intact, pas de détachement des cellules musculaires précurseurs, pas de muscles des membres.
<i>Gab1</i> -/-	Détachement et migration des cellules musculaires précurseurs diminués, certains muscles spécifiques affectés.
<i>Msx 1</i>	Serait impliqué dans la prolifération des cellules musculaires précurseurs des membres antérieurs.
<i>Six1,4(5)</i>	Exprimé dans le dermomyotome hypaxial, les cellules musculaires précurseurs en migration et les muscles des membres. Serait impliqué dans les processus de détermination et de différenciation terminale des cellules musculaires.

Ainsi, d'autres facteurs de transcription à homéodomaine ont été mis en évidence. *Lbx1*, exprimé par les cellules se détachant du myotome hypaxial est indispensable à la migration correcte de certaines de ces cellules. Une souris *Lbx1* -/- présente une absence complète de muscles squelettiques au niveau des membres alors que la formation du diaphragme et de la langue n'est pas affectée par la mutation (Buckingham *et al.*, 2003).

Mox2 également exprimé par les cellules du myotome hypaxial est indispensable à la migration et à la prolifération des cellules. En effet, une souris *Mox2* ^{-/-} présente toutes les caractéristiques d'une dystrophie musculaire (masse musculaire faible et malformation de tous les muscles squelettiques). Msx1 et les membres de la famille Six sont aussi des facteurs de transcription à homéodomaine impliqués dans la prolifération, la détermination et/ou la différenciation des cellules musculaires. c-Met, un récepteur à tyrosine kinase et son ligand SF (scatter factor) de la famille des HGF (hepatocyte growth factor), sont essentiels pour le détachement des cellules du dermomyotome. En effet, leur niveau d'expression est important pour l'amorçage du processus de migration de ces cellules. L'absence de Gab1, un autre ligand de c-Met, réduit le détachement des cellules et provoque des malformations des masses musculaires des membres et du diaphragme (Buckingham *et al.*, 2003).

3.2.2 Différenciation des précurseurs musculaires

Au terme de leur migration les précurseurs musculaires vont exprimer des facteurs de transcriptions spécifiques. Ceux sont des régulateurs myogéniques qui vont permettre aux cellules musculaires de se différencier pour former la musculature des membres.

3.2.2.1 Les protéines à domaine bHLH

3.2.2.1.1 Facteurs myogéniques (MRFs)

Des études portant sur la transcription dans le muscle squelettique ont abouti à l'isolement et à la caractérisation fonctionnelle du facteur MyoD (Davis *et al.*, 1987). D'autres travaux ont ensuite permis le clonage de plusieurs autres facteurs myogéniques : Myf5 (Braun *et al.*, 1989), Myogénine (Edmondson et Olson, 1990; Wright *et al.*, 1989) et MRF4 (également nommé Myf6) (Rhodes et Konieczny, 1989). Ces quatre facteurs régulateurs de la myogenèse (MRFs), appartenant à la famille MyoD, sont des facteurs de transcription spécifiques du muscle squelettique. Ils appartiennent à la super famille des facteurs de transcription de type basique hélice-boucle-hélice (bHLH) (Fig.32.). La région basique est responsable de leur fixation à l'ADN *via* la séquence consensus CANNTG, appelée boîte E (E-box), présente dans les promoteurs de plusieurs gènes spécifiques du muscle squelettique (Lassar *et al.*, 1989; Olson et Klein, 1994; Sabourin et Rudnicki, 2000). Le domaine bHLH permet de former des homodimères ou des hétérodimères avec d'autres facteurs à domaine

HLH, favorisant (E12 et E47) ou inhibant (Id) leur liaison à l'ADN (Murre *et al.*, 1989; Blackwell et Weintraub, 1990).

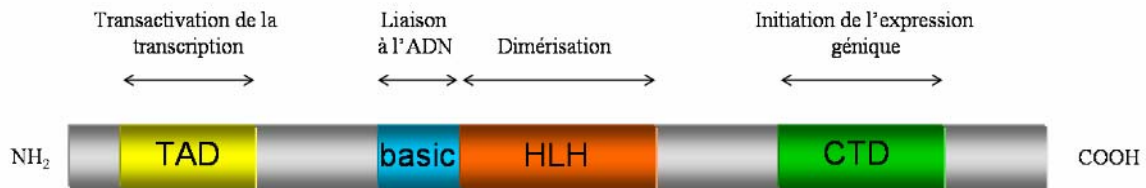


Figure 32 : Représentation schématique du facteur de transcription MyoD.

Un domaine de transactivation est présent dans la partie N-terminale de la protéine (TAD : Transactivation Amino-terminal Domain). Sa partie basique permet la fixation aux boîtes E (E-box) présentes dans les séquences promotrices de nombreux gènes. Le domaine HLH (Helix-Loop-Helix), constitué de deux hélices α , permet la dimérisation avec, entre autres, les protéines E. Le domaine C-terminal (CTD : C-terminal Domain) confère l'activité nécessaire à l'expression des gènes spécifiques du muscle squelettique.

Les MRFs ont la capacité de convertir des cellules non musculaires en cellules musculaires squelettiques *in vitro* (Weintraub *et al.*, 1991) et *in vivo* (Santerre *et al.*, 1993). Ils ont donc un rôle dans la détermination et la différenciation des précurseurs myogéniques.

3.2.2.1.2 Cinétiques d'expression

Chaque facteur myogénique a un profil d'expression spécifique au cours de la myogenèse (Fig.33.). Ce profil est différent entre les somites et les masses pré-musculaires des membres; en culture, ce profil varie également selon la lignée musculaire (Buckingham, 1992).

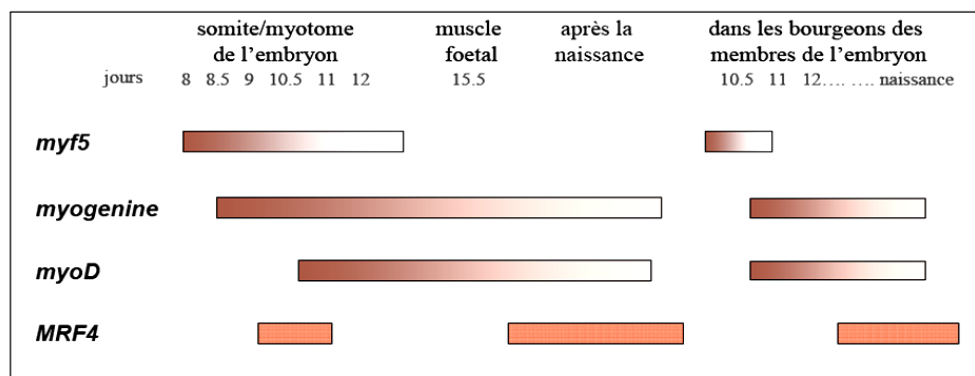


Figure 33 : Représentation schématique de l'expression des gènes myogéniques de la famille MyoD, au cours du développement musculaire chez la souris.

Ces résultats sont obtenus à partir d'hybridations *in situ* réalisées sur des embryons de souris. Ils montrent l'expression temporelle des transcrits lorsqu'ils commencent à s'accumuler dans les somites et les bourgeons des membres (d'après Buckingham, 1992).

Chez la souris, des transcrits *Myf5* sont présents très tôt dans le somite, dès 8 jours de gestation. Lorsque le somite s'est différencié, l'expression de *Myf5* se limite au myotome puis diminue progressivement pendant la myogenèse. Dans les muscles des membres, *Myf5* est également le premier facteur myogénique à être exprimé. Des transcrits *Myogénine* sont détectés peu de temps après ceux de *Myf5* dans les somites, alors que les protéines ne sont exprimées qu'à 10,5 jours de gestation, en même temps que MyoD. Dans les muscles des membres, les transcrits et les protéines Myogénine et MyoD s'accumulent en même temps. L'expression de *MRF4* est biphasique : un premier pic d'une durée de 48 heures est détectable à partir de 9,5 jours de vie embryonnaire et un second pic pendant la période de vie foetale, jusqu'à la naissance (Rawls et Olson, 1997). C'est le seul facteur myogénique à être exprimé dans les muscles adultes.

In vitro, des taux élevés de transcrits *MyoD* peuvent être détectés dans les myoblastes de certaines lignées cellulaires murines, contrairement à d'autres qui contiennent davantage de transcrits *Myf5* (Buckingham, 1992). La Myogénine est toujours exprimée dans les cellules qui se différencient, et les transcrits *MRF4* s'accumulent uniquement dans les cellules en fin de différenciation, lorsque les gènes de structure, caractéristiques des fibres matures, sont activés.

3.2.2.1.3 Activités

Plusieurs études suggèrent que la détermination et/ou la survie des précurseurs myogéniques requièrent soit MyoD soit Myf5. L'inactivation du gène *MyoD*, réalisée par Rudnicki et ses collaborateurs en 1992, n'a tout d'abord pas apporté d'indice précis sur la fonction de la protéine MyoD. En effet, la souris *MyoD*^{-/-} présente une myogenèse embryonnaire normale. L'animal adulte est viable et fertile, et, à l'exception d'un fort taux d'ARNm *Myf5* post-natal, présente un phénotype musculaire apparemment normal. De même, la souris *Myf5*^{-/-} semble avoir un phénotype musculaire normal, mais celle-ci présente une mortalité périnatale accrue due à une sévère défaillance du développement des côtes (Braun *et al.*, 1992). Les souris double KO pour *MyoD* et *Myf5* montrent une absence complète de myoblastes et de muscles squelettiques différenciés au niveau de la tête, du tronc et des membres (Kablar *et al.*, 2003) (Fig.34.), et de ce fait, meurent aussitôt après la naissance. Donc, soit les deux facteurs, MyoD et Myf5, ont des rôles compensateurs au sein de mêmes

cellules, soit ils gouvernent la différenciation de deux populations cellulaires différentes qui peuvent se compenser en l'absence de l'une ou l'autre des populations (Rawls et Olson, 1997). En fait, Myf5 semble diriger la formation des muscles à partir des cellules du dermomyotome épithélial (muscles du dos et intercostaux), et MyoD, le développement des muscles à partir des cellules migrantes (muscles des membres, du diaphragme et de la langue) (Huh *et al.*, 2005). L'expression et la localisation de Pax3 dans les embryons de souris *Myf5*^{-/-} et *MyoD*^{-/-} ne sont pas perturbées. La migration des cellules précurseurs multipotentes dans les membres n'est donc pas affectée par l'absence d'expression de ces deux facteurs. (Kablur *et al.*, 2003). Cependant, elles ne pourront pas acquiescer un devenir myogénique. MyoD et Myf5 sont donc les facteurs de détermination des myoblastes squelettiques.

Les souris dont le gène *Myogénine* a été invalidé meurent dès leur naissance avec une sévère déficience en fibres musculaires différenciées, bien que présentant un nombre normal de myoblastes déterminés (Venuti *et al.*, 1995) (Fig.34.). Son expression ne semble pas requise pour l'expression de *MyoD*, car celui-ci est normalement exprimé chez les souris *Myogénine*^{-/-}. Chez ces dernières, les masses pré-musculaires se développent de manière identique à celles des embryons sauvages; cependant, les cellules n'expriment pas les marqueurs spécifiques en même temps et en même quantité. La Myogénine est indispensable à l'induction de la différenciation terminale des cellules musculaires, en aval des actions de Myf5 et de MyoD.

MRF4 est le facteur myogénique majoritairement exprimé dans la fibre musculaire squelettique mature adulte. Les souris dont le gène *Mrf4* a été invalidé présentent des altérations de l'expression des gènes spécifiques du muscle. L'absence de MRF4 semble être compensée par une augmentation du taux de Myogénine. Néanmoins, l'expression de MRF4 est nécessaire à la différenciation terminale des cellules musculaires (Valdez *et al.*, 2000).

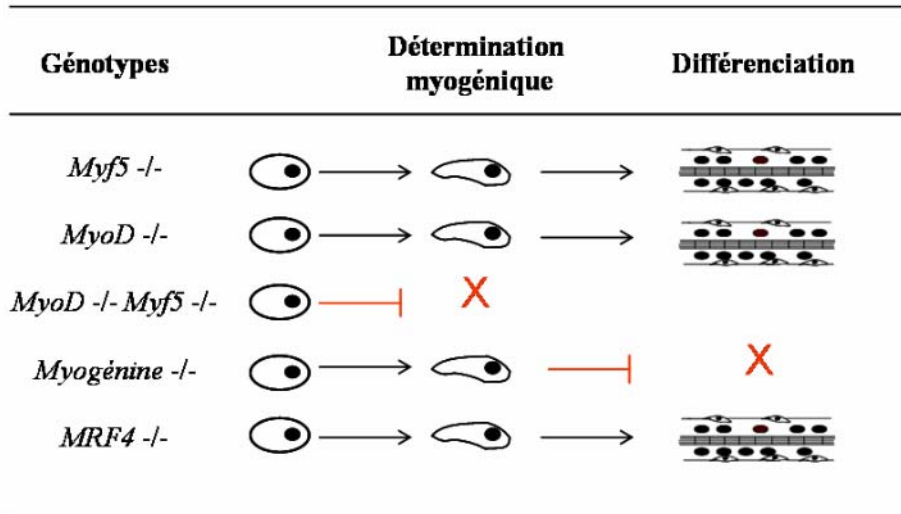


Figure 34 : Résultats des invalidations des facteurs myogéniques chez la souris. Une croix rouge indique un arrêt dans le processus de la myogenèse du fait de la mutation (d'après Guerneq, 2003).

Ce sont ces expériences d'inactivation de gènes chez la souris qui ont permis de mieux comprendre les rôles respectifs et l'importance de l'expression spatio-temporelle des facteurs myogéniques dans la formation des muscles. La figure 35 résume les différentes étapes de la formation d'une fibre musculaire et l'expression temporelle des facteurs myogéniques.

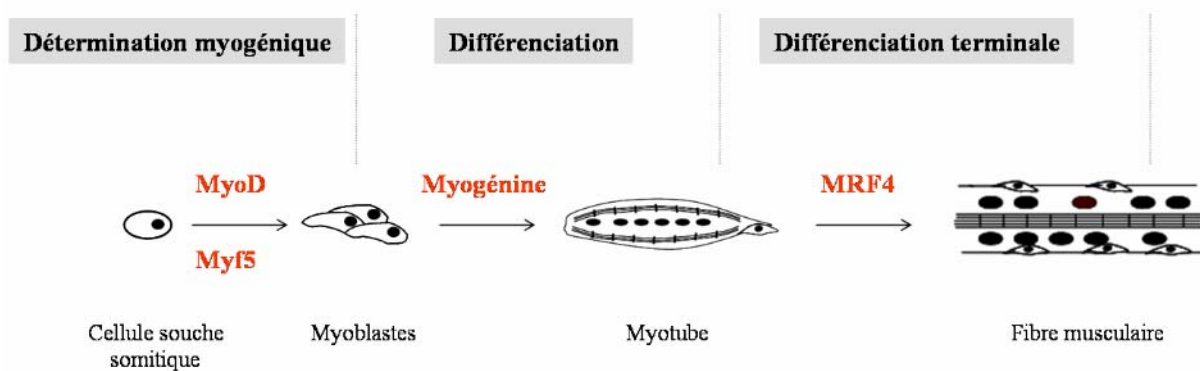


Figure 35 : Expression des facteurs myogéniques de la famille MyoD au cours de la myogenèse. MyoD et Myf5 induisent la détermination des myoblastes alors que la Myogénine et MRF4 sont nécessaires à la différenciation.

3.2.2.2 Les protéines à domaine MADS

La famille des protéines à domaine MADS regroupe divers facteurs de transcription. Elle a été définie sur la base d'homologies de séquences entre des protéines issues d'organismes eucaryotes, dont la levure, les Insectes et les Mammifères (Shore et Sharrocks, 1995). Le motif conservé, appelé domaine MADS (MCM1-Agamous-Deficiens-SRF), est composé de 56 acides aminés - dont 9 sont strictement conservés chez tous les membres de la famille - et de deux régions fonctionnelles - l'extrémité N-terminale qui constitue le déterminant majeur de spécificité de liaison à l'ADN et l'extrémité C-terminale qui est nécessaire à la dimérisation.

3.2.2.2.1 Facteur de réponse au sérum (SRF)

C'est l'un des membres les plus étudiés de la famille de protéines à domaine MADS. D'une masse moléculaire de 67 KDa, ce facteur de transcription se fixe sous forme d'homodimère sur une séquence de 10 nucléotides, CC(AT)₆GG, appelée boîte CArG. Les fonctions de la protéine sont multiples puisqu'il existe une trentaine de sites CArG fonctionnels dans les promoteurs de différents gènes dont celui du gène de l'actine α (Chai et Tarnawski, 2002). Son inactivation génique n'apporte que peu d'indices quant à ses différentes fonctions. En effet, les embryons *SRF*^{-/-} se développent normalement jusqu'à E6,5 puis meurent *in utero* lors de la gastrulation, essentiellement par absence de formation du mésoderme (Arsenian *et al.*, 1998). Ainsi, des études fonctionnelles menées sur la protéine ont été en grande partie réalisées *ex vivo* dans des lignées cellulaires. Elles ont permis de distinguer deux fonctions principales de la protéine SRF : l'induction de gènes impliqués dans la prolifération et le cycle cellulaire et la régulation des processus de différenciation (Chai et Tarnawski, 2002; Miano, 2003).

La protéine SRF est exprimée de façon constitutive dans les cellules musculaires en culture. Son niveau d'expression et sa localisation nucléaire sont les mêmes dans les cellules en prolifération et dans les cellules différenciées. Une étude de Vandromme et ses collaborateurs en 1992 démontre que l'inactivation de la protéine SRF, par l'injection d'anticorps anti-SRF, inhibe la différenciation de cellules musculaires en culture, en empêchant la transition myoblaste-myotube. Ceci s'accompagne de l'inhibition de l'expression de la Troponine T, marqueur tardif de la différenciation, et de la Myogénine, témoignant d'un rôle précoce de la protéine SRF dans le programme de différenciation myogénique. Il est

intéressant de noter qu'à l'inverse du promoteur de la Troponine T qui est effectivement une cible directe de la protéine SRF, aucune séquence CArG fonctionnelle n'a été identifiée dans les séquences régulatrices du gène *Myogénine*. L'inhibition de l'expression de la Myogénine par inactivation de la protéine SRF témoigne d'un rôle général de cette dernière dans le processus de différenciation. Des expériences similaires ont démontré que l'inactivation de la protéine SRF conduit à une perte de l'expression de la protéine MyoD sans toutefois altérer l'expression du facteur Myf5 (Carnac *et al.*, 1998; Wei *et al.*, 1998).

3.2.2.2 Famille des Myocyte Enhancer Factors 2 (MEF2)

En 1989, MEF2 est initialement décrit comme un facteur nucléaire présent dans les cellules musculaires différenciées, mais absent des cellules en prolifération et des cellules non musculaires. Cette protéine se lie à une séquence consensus riche en A/T (CTA(A/T)₄TAG) au niveau des promoteurs des gènes musculaires. Elle est décrite comme nécessaire à l'expression de la Créatine Kinase Musculaire (CMK) et de la chaîne légère de la myosine de rat (Gossett *et al.*, 1989). Depuis, une famille de protéines MEF2 a été caractérisée après clonage de leurs ADNc. Chez les Vertébrés, elle se compose de quatre gènes *mef2* appelés *mef2a* (*myocyte enhancer factor 2a*), *mef2b*, *mef2c* et *mef2d*, localisés, chez la souris, sur des chromosomes différents (Morisaki *et al.*, 1997). Chaque isoforme protéique possède, outre un domaine MADS, un domaine MEF2 amino-terminal de 29 acides aminés. Ce dernier représente une extension du domaine MADS constituant une région qui assure essentiellement la liaison de la protéine à l'ADN, mais aussi son interaction avec des partenaires. Les membres de la famille MEF2 peuvent s'homo- ou s'hétérodimériser. Le site de liaison MEF2 présent dans de nombreux promoteurs peut aussi lier des facteurs autres que ceux de la famille MEF2, créant ainsi des mécanismes de compétition qui participent à la régulation de certains gènes.

Il existe une régulation bidirectionnelle entre les MRFs et les MEF2. En effet, une activité MEF2 est induite dans des cellules non musculaires transfectées par la Myogénine (Cserjesi *et al.*, 1994) ou MyoD (Yu *et al.*, 1992), ce qui suggère que les MRFs peuvent activer MEF2. Dès 1992, plusieurs études démontrent l'existence dans le promoteur du gène *Myogénine*, d'un site MEF2 fonctionnel nécessaire à l'expression du gène dans des lignées cellulaires (Edmondson *et al.*, 1992; Buchberger *et al.*, 1994). En 1995, deux études respectivement de Naidu et Black révèlent l'existence, dans le promoteur du gène *MRF4*, d'un site MEF2 nécessaire à la régulation positive du gène. En effet, sa mutation entraîne une forte

réduction d'activité du promoteur lors de la différenciation de cellules myoblastiques murines de la lignée C2C12 et de myoblastes primaires de poulet (Black *et al.*, 1995). Outre un site MEF2, le promoteur du gène *MRF4* contient aussi une boîte E fonctionnelle (Naidu *et al.*, 1995). Des études de cotransfection démontrent l'activation synergique du promoteur de *MRF4* en présence des deux facteurs MEF2 et Myogénine. Tandis qu'une double mutation de leurs sites fixateurs entraîne une perte totale d'activité, toute simple mutation de l'une ou l'autre des boîtes E ou MEF2 n'altère pas la synergie entre les deux facteurs. Ce résultat démontre l'existence d'une coopération entre ces deux facteurs sur l'un ou l'autre des deux sites et suggère leur interaction (Rogerson *et al.*, 2002). Cependant, une telle interaction entre les MRFs et les MEF2 n'est pas toujours nécessaire. En effet, une étude récente montre que l'expression de *mef2c*, le premier à être exprimé dans le muscle squelettique, est directement activée par les facteurs bHLH qui se fixent sur les boîtes E de sa région promotrice, et que le site riche en A/T aussi présent au niveau de son promoteur ne lie aucun membre de la famille MEF2 pour autoréguler son expression (Dodou *et al.*, 2003).

Les protéines MEF2 sont impliquées dans le contrôle des étapes tardives de la différenciation des cellules musculaires, puisque d'une part, elles ne peuvent l'induire seule (Molkentin *et al.*, 1995), et d'autre part, parce que leur expression est induite par les facteurs myogéniques (Wang, Valdez *et al.*, 2001; Dodou *et al.*, 2003).

3.3 Effecteurs de la myogenèse

Au début du processus myogénique, l'organisation des somites en plusieurs compartiments dépend des structures avoisinantes, en l'occurrence du tube neural, de la notochorde et de l'ectoderme dorsal (Buckingham, 2001). En effet, ces tissus sécrètent ou exposent à leurs surfaces des facteurs capables d'orienter la destinée des cellules somitiques.

De nombreux signaux extracellulaires régulent la transcription des facteurs myogéniques et ainsi modulent les différents stades du développement du muscle squelettique, de la détermination des précurseurs myogéniques jusqu'à leur différenciation et la formation des fibres musculaires. Ces facteurs, exprimés dans les régions axiale et latérale de l'embryon en développement, sont importants pour la formation des somites et la détermination des différentes lignées de cellules musculaires, à savoir, les lignées myogéniques épaxiale et hypaxiale. Ils comprennent entre autres les Wnts (Wingless), les IGFs (Insulin-like Growth Factors), Notch, les FGFs (Fibroblast Growth Factors) et les TGF β s (Transforming Growth Factors- β) tels que les BMPs (Bone Morphogenetic Proteins). Tous ces facteurs régulent la détermination et la différenciation myogénique (Fig.36.).

3.3.1 Les inducteurs

3.3.1.1 La famille des Insulin-like growth factors (IGFs)

Les IGFs appartiennent à une famille de petits peptides caractérisés par une structure très proche de la pro-insuline humaine. Les plus importants sont IGF I, II et l'insuline. Ce sont des régulateurs importants du programme de différenciation myogénique. L'expression des IGFs et de leurs récepteurs, est détectée très tôt au cours de l'embryogenèse, au niveau du tube neural et du mésoderme paraxial. Ces facteurs stimulent à la fois la prolifération en augmentant l'expression des acteurs du cycle cellulaire et la différenciation des myoblastes en

culture (Kocamis *et al.*, 2002). Par la suite, ils activent les facteurs de transcription myogénique provoquant l'expression des gènes structuraux nécessaires à la formation des myotubes du muscle squelettique (Zorzano *et al.*, 2003). Une étude réalisée *in vitro* sur des somites de poulets en 2000 par Pirskanen et ses collaborateurs démontre que les trois principaux IGFs sont capables d'induire la myogenèse, ceci en synergie avec d'autres facteurs tels que les β FGF, le TGF- β 1 et Shh (Sonic Hedgehog)

Ces facteurs peuvent également induire une hypertrophie des myotubes lors du développement du muscle squelettique (Zorzano *et al.*, 2003). En effet, l'ajout d'IGF-I à des myoblastes en culture provoque une augmentation de leur activité mitotique, suivie de l'induction de leur programme de différenciation (Semsarian *et al.*, 1999). Cet effet passe par une altération des niveaux d'expression de Myf5, p21, Myogénine, MRF4, et de certaines cyclines. Les souris transgéniques surexprimant IGF-I présentent une hypertrophie musculaire (Musaro *et al.*, 1999), alors que la délétion des gènes codant IGF-I, IGF-II et les récepteurs IGF, conduit à une diminution de la masse musculaire associée à des troubles du développement (Kocamis *et al.*, 2002).

3.3.1.2 La famille Wingless (WNTs)

Les membres de la famille Wnt sont des glycoprotéines sécrétées, impliquées dans de nombreux processus développementaux. Aujourd'hui, 19 membres de cette famille sont identifiés chez l'Homme (Miller, 2002). Ce sont des homologues de la protéine Wingless, premièrement identifiée chez la drosophile (Rijsewijk *et al.*, 1987).

Les récepteurs des Wnts sont les protéines Frizzled (10 membres chez l'Homme) qui comptent sept domaines transmembranaires. La fixation des Wnts à leurs récepteurs entraîne l'activation d'une cascade de signalisation conduisant à l'expression de gènes spécifiques du muscle (Cossu et Borello, 1999). Ainsi, il a été démontré que dans des explants de mésoderme présomitique de souris, les structures axiales (tube neural) exprimant Wnt-1, 4 et 5a, activent préférentiellement dans le dermomyotome l'expression de *Myf5* tandis que l'ectoderme dorsal exprimant Wnt-7a, active préférentiellement *MyoD* (Cossu *et al.*, 1996).

Des expériences d'inactivation chez le poulet ont montré que les Wnts agissent en synergie avec Shh. En effet, ce dernier sécrété par la notochorde rend les cellules progénitrices musculaires compétentes pour l'activation des gènes myogéniques en présence

des Wnts. Cependant, il semble que Shh soit nécessaire seulement de manière transitoire puisqu'il n'est plus requis à des étapes plus tardives de la somitogenèse (Ridgeway *et al.*, 2000).

Des protéines antagonistes des WNTs ont été mises en évidence. Par exemple, la protéine Cerberus se fixe directement aux WNTs pour inhiber leur fixation à leurs récepteurs et permet ainsi la formation des structures de la tête (Piccolo *et al.*, 1999). La protéine WIF-1 (WNT inhibitory factor-1) se fixe elle aussi aux WNTs pour inhiber leur fixation à leurs récepteurs au niveau des somites et permet leur compartimentation. En effet, des expériences d'ARN interférence réalisées afin de supprimer l'expression de cette protéine au niveau des somites d'embryons de xénope ont entraîné la formation d'un second axe antéro-postérieur et une désorganisation des somites (Kawano et Kypta, 2003). WIF-1 est une protéine sécrétée possédant cinq domaines EGF. Une recherche *in silico* dans les bases de données moléculaires et une analyse de la séquence m'ont permis d'identifier que deux de ses domaines EGF chez l'Homme et la souris (le 3^{ème} et le 5^{ème}) sont potentiellement *O*-fucosylables. Aucune étude n'a encore porté sur le rôle de sa glycosylation dans ses interactions avec les WNTs.

3.3.2 Les inhibiteurs

3.3.2.1 La famille des Transforming growth factors β (TGF β)

Plus de quarante membres ont été identifiés dans cette superfamille qui peut être subdivisée en deux sous-familles majeures : les TGF- β de type Activine/Nodal et les BMP/GDF/MIS (Bone Morphogenetic Protein/Growth and Differentiation Factor/Muellerian inhibiting substance). Ces différents facteurs sont regroupés sur la base d'homologies de séquences et sur la spécificité de la voie de signalisation qu'ils activent. Ces ligands partagent un grand nombre de récepteurs dont les arrangements induiront différentes voies de signalisation. La superfamille des TGF- β joue un rôle très important dans les stades précoces de l'embryogenèse, et dans le maintien de l'homéostasie des tissus adultes, en régulant la croissance et la différenciation cellulaire, ainsi que l'apoptose (Shi et Massague, 2003).

Les TGF β sont des inhibiteurs de la différenciation myogénique. Ils bloquent de nombreuses étapes dont la fusion des myoblastes et l'accumulation de protéines spécifiques du

muscle. Le traitement de lignées cellulaires de muscles squelettiques et de cellules primaires musculaires par les TGF β inhibe, de façon dose-dépendante, l'expression de tous les marqueurs principaux de la différenciation myogénique, mais stimule la prolifération des cellules (Florini *et al.*, 1991).

La famille des BMPs comprend plus de trente membres. Ce sont des protéines sécrétées, impliquées dans divers processus du développement embryonnaire des Vertébrés, notamment dans la formation des tissus musculaires et osseux. Le facteur le plus étudié est BMP4. Il est exprimé par le tube neural et inhibe le développement du myotome en réprimant des facteurs myogéniques et en maintenant l'expression du gène *Pax3* (Pourquie *et al.*, 1996).

3.3.2.2 La famille des Fibroblast growth factors (FGFs)

Les FGFs sont des facteurs solubles séquestrés dans un premier temps par des protéoglycanes à héparane-sulfate dans la matrice extracellulaire, puis libérés pour une utilisation locale. Le traitement des myoblastes en culture avec ces facteurs stimule leur prolifération et inhibe leur différenciation. *In vivo*, ces molécules contribuent à la formation du myotome dorsal en stimulant la prolifération et la survie des cellules (Zorzano *et al.*, 2003). Les FGFs sont exprimés très tôt dans la partie dorsale du tube neural, au moment où les cellules somitiques acquièrent leur caractère myogénique (Stern *et al.*, 1997). Ils sont indispensables à l'activation de la myogenèse dans le mésoderme paraxial. En effet, la signalisation par les récepteurs aux FGFs induit la prolifération des myoblastes et leur migration dans les bourgeons des membres (Buckingham *et al.*, 2003).

Les FGF β sont les membres de cette famille les plus largement distribués et les plus compétents pour l'inhibition de la différenciation des cellules musculaires (Florini *et al.*, 1991). *In vitro*, ce sont de puissants mitogènes pour les myoblastes, ce qui a pour conséquence de bloquer le démarrage de la différenciation. Ces facteurs sont impliqués dans le contrôle négatif de l'expression des transcrits *Myogénine* et *MyoD* (Iwahori *et al.*, 2004).

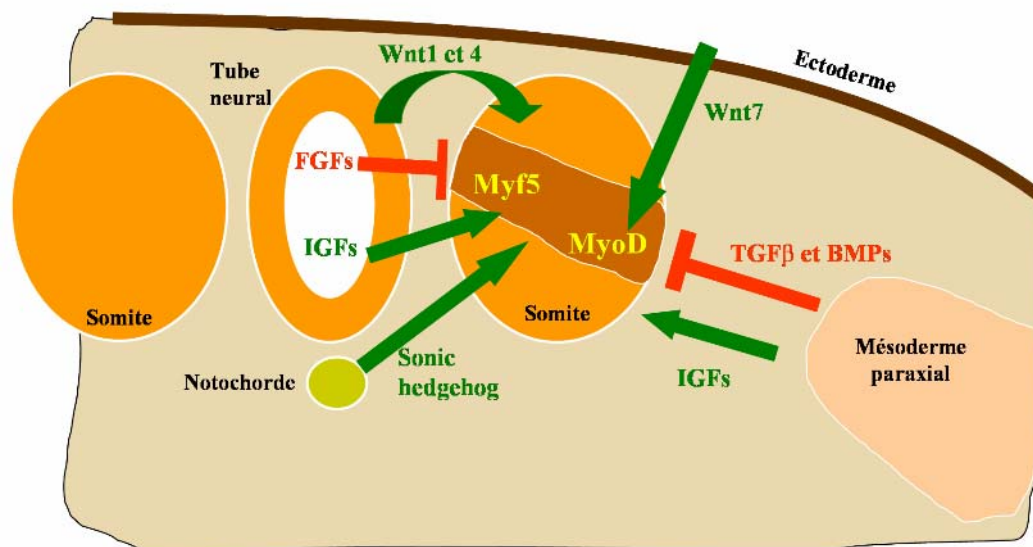


Figure 36 : Schéma récapitulatif de l'effet inducteur (vert) ou inhibiteur (rouge) des différents effecteurs de la myogenèse, secrétés par les structures environnant les somites.

3.3.2.3 Notch et ses ligands

Les récepteurs Notch ainsi que leurs ligands sont des protéines transmembranaires présentant dans leurs parties extra-cellulaires de nombreux domaines EGF. Ces derniers, possédant de nombreux sites de *O*-fucosylation, sont nécessaires à l'interaction ligand-récepteur.

Au sein d'un groupe de cellules équivalentes, certaines portant le récepteur Notch reçoivent un signal d'inhibition latérale de la part d'une ou de plusieurs cellules voisines présentant à leur surface les ligands de Notch (Delta ou Serrate/Jagged). Elles sont alors restreintes dans leur capacité à se différencier, alors que les cellules émettrices du signal poursuivent leur différenciation (Schweisguth *et al.*, 1996). Cette voie de signalisation a été conservée au cours de l'évolution chez les Animaux, cependant bien que de plus en plus de gènes impliqués dans cette voie soient connus, les mécanismes moléculaires induisant l'activation ou la répression de gènes cibles restent partiellement incompris.

La liaison du ligand au récepteur Notch induit un clivage protéolytique intracellulaire du récepteur. Le domaine intracytoplasmique de Notch migre alors dans le noyau où il va stimuler ou inhiber la transcription de gènes cibles (Jarriault *et al.*, 1995; Schweisguth *et al.*, 1996). Ce mode de signalisation a d'abord été montré sur des formes constitutivement actives

de Notch délété des répétitions EGF (Kopan *et al.*, 1996), puis sur le récepteur entier chez les Mammifères (Schroeter *et al.*, 1998) et chez la drosophile (Kidd *et al.*, 1998).

Parmi les 4 récepteurs Notch présents chez les Mammifères, Notch 1, 2 et 4 sont plus particulièrement impliqués dans la myogenèse. C'est sur le récepteur Notch1 que porte la plupart des études. Celui-ci régule la myogenèse à trois niveaux :

> Une première activation de la voie de signalisation par Notch est nécessaire à la mise en place des somites, au moment de la transition entre mésoderme pré-somitique et mésoderme somitique. Il semble que celle-ci joue un rôle critique dans la segmentation des somites, en déterminant probablement par inhibition latérale au sein de groupes de cellules équipotentes, le nombre de cellules qui auront un destin somitique. En effet, chez une souris KO pour le gène *Notch1*, la formation des somites est plus tardive et irrégulière (Conlon *et al.*, 1995).

> Une seconde activation serait nécessaire au sein du dermomyotome pour empêcher certaines cellules d'acquies un devenir musculaire (Nofziger *et al.*, 1999). Après interaction du récepteur avec un de ses ligands, le domaine intracytoplasmique de Notch1 migre dans le noyau de la cellule où il est capable d'inhiber par contact direct le facteur de transcription E47 (un facteur de type basique HLH) (Fig.37.1). Ce dernier ne peut donc plus former d'hétérodimères avec MyoD et ainsi activer l'expression de gènes spécifiques du muscle (Luo *et al.*, 2005).

> La troisième activation du récepteur Notch se produit dans une population de pré-myoblastes afin d'empêcher leur différenciation avant d'avoir atteint les bourgeons des membres. Dans ce cas, le domaine intracytoplasmique de Notch migre dans le noyau où il va s'associer avec un facteur de régulation négatif de la transcription, la protéine Su(H) chez la drosophile ou CBF1-RBP chez les Mammifères. Ceci permet la transcription des facteurs HES1 et HES5 (Hairy/Enhancer of Split genes) qui était jusqu'alors réprimée (Mumm et Kopan, 2000). Ces derniers vont inhiber de façon indirecte la transcription de gènes myogéniques en bloquant le facteur de transcription MyoD (Fig.37.2). De plus, le domaine intracytoplasmique de Notch est capable de bloquer, par contact direct, MyoD ainsi que le facteur MEF2c de la famille des protéines à domaine MADS (Wilson-Rawls *et al.*, 1999) (Fig.37.3).

Les voies de signalisation par les récepteurs Notch fonctionnent comme une horloge biologique qui se déclenche pour déterminer le devenir de certaines cellules dans une population, et notamment celui des cellules à l'origine des muscles squelettiques. Elles sont également impliquées dans la prolifération et la différenciation des cellules souches musculaires dans le cas de lésions (Luo *et al.*, 2005). Les mécanismes de régulation de ces voies de signalisation sont extrêmement complexes et tous les acteurs moléculaires ne sont pas encore déterminés. On sait cependant que des protéines cytoplasmiques comme Numb, ou secrétées comme les WNTs, sont des antagonistes de Notch alors que la protéine Deltex agit comme un régulateur positif (Fig.37.4) (Luo *et al.*, 2005).

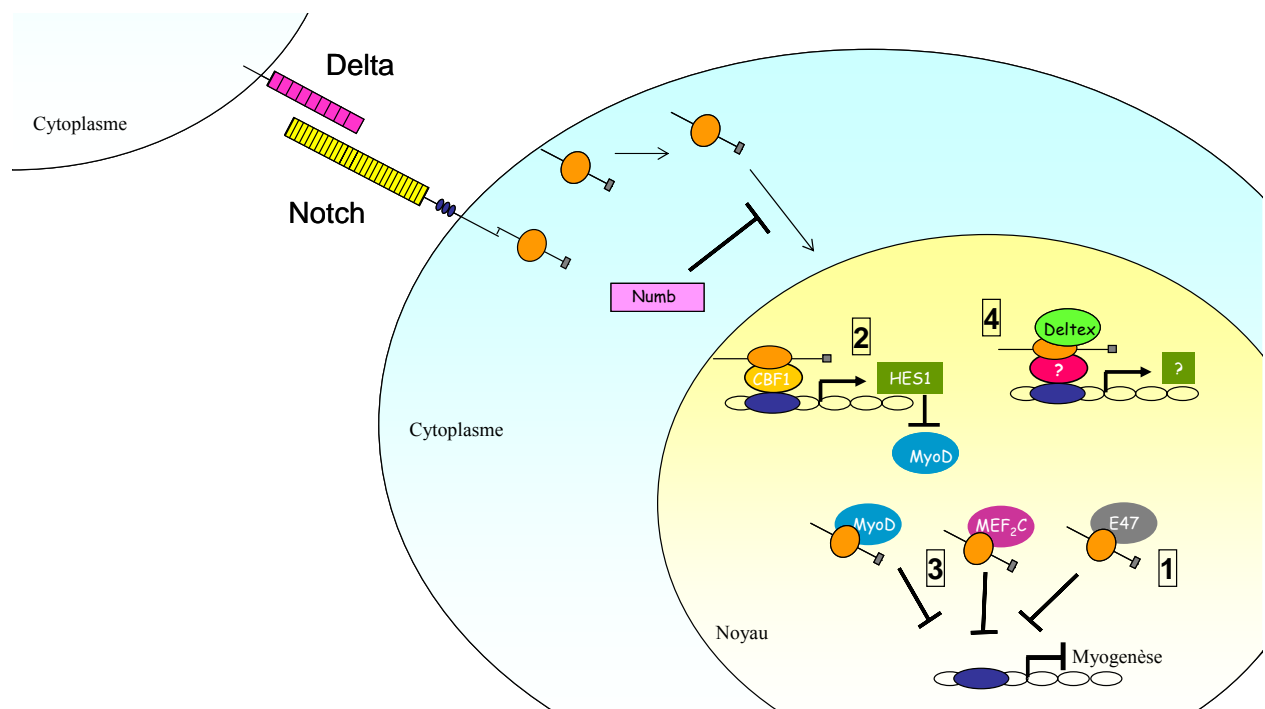


Figure 37 : Représentation schématique des voies de signalisation du récepteur Notch contrôlant la myogenèse.

3.4 La myogenèse *in vitro*

Pour faciliter les recherches sur la myogenèse, divers systèmes *in vitro* ont été développés. Ce sont soit des cultures primaires de cellules musculaires squelettiques provenant essentiellement de muscles pectoraux d'embryon de poulet ou de caille et de muscles des membres postérieurs de rat ou de souris nouveaux-nés, soit des cultures de lignées cellulaires myogéniques stabilisées. Ces systèmes de culture cellulaire reproduisent les diverses phases de la myogenèse, détermination, prolifération, différenciation, et constituent donc de bons modèles d'étude de celle-ci. En particulier, il a été possible de définir les conditions optimales afin de faire différencier des myoblastes non encore spécialisés du point de vue morphologique, en myotubes polynucléés, et ceci afin d'analyser les modifications intervenant au cours de la différenciation myogénique. Dans les années 60, un important progrès fut accompli dans l'étude de la différenciation cellulaire *in vitro*, par la mise en place de cultures clonales de cellules embryonnaires non spécialisées. Leur développement a rendu possible l'étude des conditions déterminant la prolifération et la différenciation des cellules myogéniques. L'identification des facteurs moléculaires déterminant ces fonctions a permis d'accroître les connaissances du processus myogénique *in vitro* et *in vivo*.

L'étude des mécanismes qui régulent le processus myogénique chez les Mammifères est réalisée principalement sur deux types de lignées cellulaires immortalisées, la lignée C2C12 et la lignée L6.

> La lignée C2C12 est un sous-clone dérivé de la lignée originelle C2 obtenue par culture primaire de myoblastes isolés à partir de la patte postérieure d'une souris (Yaffe, 1968). Parmi les cellules adhérentes obtenues, le clone C2C12 spontanément immortalisé a été sélectionné. La lignée C2C12 est caractérisée par sa capacité à se différencier rapidement et à former de grands myotubes.

> La lignée L6 dérive de myoblastes obtenus à partir de la patte postérieure d'un rat nouveau-né, stabilisés par traitement avec un agent carcinogène chimique (Yaffe et Saxel, 1977). Les cellules de cette lignée constituent un bon modèle pour l'étude de la myogenèse. Elles ont une forte capacité de différenciation dans des milieux de culture contenant peu de sérum. Néanmoins, la multiplication des passages diminue la capacité de fusion de ces cellules, rendant nécessaire le reclonage périodique de la lignée.

Résultats-Discussion

Cette nouvelle partie concerne les résultats obtenus au cours de ma thèse. Elle est scindée en trois chapitres correspondant respectivement à deux articles et aux derniers résultats obtenus. Les deux premiers chapitres constituent le volet fondamental du travail dont le but est d'étudier l'évolution des gènes de O-fucosyltransférases, de caractériser les gènes bovins Pofut1 et Pofut2, et d'analyser plus particulièrement la structure de l'enzyme Pofut1, notamment le rôle de ses N-glycanes. Le troisième chapitre, qui ne s'appuie pas pour l'instant sur un article, est consacré à l'étude de l'implication de l'enzyme Pofut1 au cours de la différenciation des cellules myoblastiques en culture. Pour les deux premiers chapitres, après un résumé succinct des résultats présentés dans l'article correspondant, des résultats complémentaires et non publiés sont exposés.

1 Evolution des gènes de *O*- fucosyltransférases. Structure et expression des gènes *Pofut1* et *Pofut2* chez *Bos taurus*.

1.1	PUBLICATION 1	125
1.2	RESULTATS COMPLEMENTAIRES	129
1.2.1	<i>Expression tissulaire des variants transcriptionnels</i>	129
1.2.2	<i>Production d'anticorps anti-Pofut bovins</i>	134
1.2.3	<i>Activités des protéines codées par les différents variants</i>	138
1.2.4	<i>Evolution des gènes de O-fucosyltransférases</i>	143
1.2.5	<i>Synthèse et conclusion</i>	152

Afin d'examiner le lien entre la *O*-fucosylation et les processus de différenciation de la cellule musculaire, il est avant tout nécessaire de caractériser les gènes et les enzymes responsables de cette modification post-traductionnelle chez notre modèle d'étude, l'espèce *Bos taurus*.

Lorsque ce travail a débuté, seule la structure génomique complète de *Pofut1* chez l'Homme et la souris était disponible (Wang *et al.*, 2001). Un seul variant transcriptionnel codant l'enzyme Pofut1 active était décrit. Les recherches sur *Pofut2* débutant à peine, ni la structure du gène ni celles d'éventuels variants transcriptionnels n'étaient décrites et l'activité de l'enzyme Pofut2 restait encore inconnue.

La caractérisation fonctionnelle de ces gènes chez le bovin a nécessité dans un premier temps la détermination des structures génomiques des gènes puis l'analyse de l'expression des transcrits correspondants. Ainsi, nous avons identifié pour la première fois, cinq variants transcriptionnels pour chacun des deux gènes, différemment exprimés au sein de treize tissus bovins. Des tests *in vitro* ont permis de vérifier les fonctions enzymatiques des protéines bovines Pofut1 et Pofut2 codées par ces différents variants.

Les comparaisons de séquences des Pofut chez différents organismes et une étude phylogénétique ont permis de confirmer que pour chaque espèce étudiée, une seule copie de chacun des gènes de *O*-fucosyltransférase est présente. Les organisations génomiques sont conservées chez les Vertébrés. De plus, les positions conservées de certains introns de *Pofut1* et de *Pofut2* permettent de conclure que très probablement l'ancêtre des Bilatériens possédait déjà des gènes morcelés.

Ces travaux sont présentés de manière détaillée dans l'article ci-après, publié dans la revue *Glycobiology* en 2006.

1.1 Publication 1

Molecular evolution of protein *O*-fucosyltransferase genes and splice variants

**Loriol, C., Dupuy, F., Rampal, R., Dlugosz, M.A., Haltiwanger, R.S., Maftah, A., and
Germot, A.**

Glycobiology, 2006, vol 16, 736-747

Tables in supplementary data

Table I. Primers used for RT-PCR analyses of bovine *Pofut1* and *Pofut2* transcripts.

Primers	Sequences
<i>Pofut1</i> ADNc/5'	5' -TTCCCTCCCAGTCTCCGCGC-3'
<i>Pofut1</i> ADNc/5'Nested	5' -GGCGCGCTCTGGCACTGGACC-3'
<i>Pofut1</i> ADNc-V1a/3'	5' -AGTATGGCCTAGGAAGTGGCG-3'
<i>Pofut1</i> ADNc-V1a/3'Nested	5' -AGTGGCTCAGGCCAGGCAGTG-3'
<i>Pofut1</i> ADNc-V1b/3'	5' -ATAAGTAGCTACCATTACTCTCTTATTAG-3'
<i>Pofut1</i> ADNc-V1b/3'Nested	5' -ATGGCCAACATCCAGAAGAGCGAC-3'
<i>Pofut2</i> ADNc/5'	5' -GCGCCCGGGGCCATGGCG-3'
<i>Pofut2</i> ADNc/5'Nested	5' -ATGGCGGCGCTTGCCGTC-3'
<i>Pofut2</i> ADNc/3'	5' -TCAGTGGTCACCCCAGGTGTTG-3'
<i>Pofut2</i> ADNc/3'Nested	5' -TGAGGCTGGACTGACCCTGC-3'

Table II. Primers used to amplify and sequence bovine *Pofut1* and *Pofut2* genes.

<i>Pofut1</i> Primers	Sequences
<i>BosPofut1</i> /5'end	5' -ATGCATGGGCAGTAGAGC-3'
Intron 1/5'	5' -GCGCTCTGGCACTGGACCCG-3'
Intron 2/5'	5' -CGTTTTGGGAACCAGGCTGATCATTTC-3'
Intron 2/3'	5' -GGTGCCAGCTTCTCCATGAAGTCCTCC-3'
Intron 3/5'	5' -TCCTACCAGAAGTACTTCAAGCTGGAG-3'
Intron 3/3'	5' -ATCCACTGGTCTTTGTAGGAGGCACTG-3'
Intron 4/5'	5' -GGCCCATTTTGGGATCAGTTTCATGTG-3'
Intron 4/3'	5' -GGATCTGGATCAGCTCACAC-3'
Intron 5/5'	5' -CATTCCCCACTTTCACCTCAA-3'
Intron 5/3'	5' -CCCCTGTCCCTACCATCTCGTCTGACC-3'
Intron 6/5'	5' -CAGAGGAACATCCAGTGC TTG-3'
Intron 6/3'	5' -TTTGAAGAGCTGCTGGATCTC-3'
Intron 7/5'	5' -ACGCGTGTGCCATGCTGAAAG-3'
Intron 7/3'	5' -CCTGTCCATGCCAAAGAAGGAGG-3'
Intron 8/5'	5' -GAAGGTGGTGAGTCTGAAGCC-3'
Intron 8/3'	5' -CGAGCAGCAAGGACTACTCTG-3'
<i>BosPofut1</i> /3'end	5' -AGCCGAGTTTTGCTGTTCCGAAG-3'
<i>Pofut2</i> Primers	Sequences
<i>BosPofut2</i> /5'end	5' -GCGATGAACTGCTCGTACTC-3'
Intron 1/5'	5' -TTGCCGTCGTCTTCC TGCTG-3'
Intron 2/5'	5' -CCTTCTGTACGATGTCAATCC-3'
Intron 2/3'	5' -GTAATACTCGTGTCTTGTCC TG-3'
Intron 3/5'	5' -CAGGACAAGCACGAGTATTAC-3'
Intron 3/3'	5' -TTCGTAACCCCAAACCATCC-3'
Intron 4/5'	5' -TGGTTTTGGGGTTACGAAGAG-3'
Intron 4/3'	5' -TACTCCTTCCCGCCGTAGTGG-3'
Intron 5/5'	5' -GAGGAGGACTGGACCCGGATG-3'
Intron 5/3'	5' -CCCAGCGAGGAGCCCAGCTTG-3'
Intron 6/5'	5' -AGCATGGTGTTCGCCAGGCAC-3'
Intron 6/3'	5' -ACAGCATCTGTGGCCACGAAC-3'
Intron 7/5'	5' -CAAGCTGGGCTCCTCGCTGGG-3'
Intron 7/3'	5' -CGGGCGTGCGCACAGATCCAC-3'
Intron 8/5'	5' -AGGAGTTGCGGAGGCTGCTGC-3'
Intron 8/3'	5' -AGTATGCAATCCTCCAATGAGTGG-3'
<i>BosPofut2</i> /3'end	5' -TCATTGGAGGATTGCATACTG-3'

1.2 Résultats complémentaires

1.2.1 Expression tissulaire des variants transcriptionnels

Une recherche dans les banques de données moléculaires disponibles par Internet a permis de mettre en évidence des EST bovins correspondant aux gènes *Pofut1* et *Pofut2*. Grâce à la sélection d'amorces spécifiques de part et d'autre des CDS de ces EST, cinq variants transcriptionnels pour chacun des deux gènes ont pu être caractérisés et leurs expressions au sein de treize tissus bovins adultes, analysées (cf. Publication 1, figure 5A et B).

Nous avons en effet choisi d'utiliser un panel assez large de tissus. Les mêmes tissus ont été prélevés sur trois bovins Charolais adultes âgés de 19 mois du centre INRA de Clermont-Ferrand-Theix, disséqués et congelés sur place par nos soins. Nous avons sélectionné cinq muscles squelettiques ayant des métabolismes énergétiques différents : lent/oxidatif, c'est-à-dire équipés de fibres adaptées à un effort d'intensité sous-maximale et de longue durée (*Rectus abdominis*, Diaphragme et Onglet) et rapide/glycolytique, c'est-à-dire équipés pour un effort intense et bref (*Semitendinosus* et *Longissimus thoraci*).

Les analyses ont été effectuées de façon indépendante sur ces trois animaux. Les résultats étant reproductibles, nous ne présentons que ceux obtenus sur un individu.

Quelque soit le tissu analysé et contrairement à leurs homologues humains et murins (Wang *et al.*, 2001, Shi et Stanley, 2003), les gènes *Pofut1* et *Pofut2* semblent faiblement exprimés chez le bovin. Ainsi, des PCR classiques sur des ADNc obtenus après transcription inverse d'ARNm extraits des différents tissus bovins n'ont jamais permis d'obtenir d'amplification. Il a donc fallu systématiquement effectuer des PCR nichées avec deux couples d'amorces.

Les tailles extrêmement proches des variants correspondant au gène *Pofut2*, n'ont pas permis de les séparer sur un gel d'agarose (Fig.38.), contrairement à ceux du gène *Pofut1* présentés dans la figure 5A de l'article.

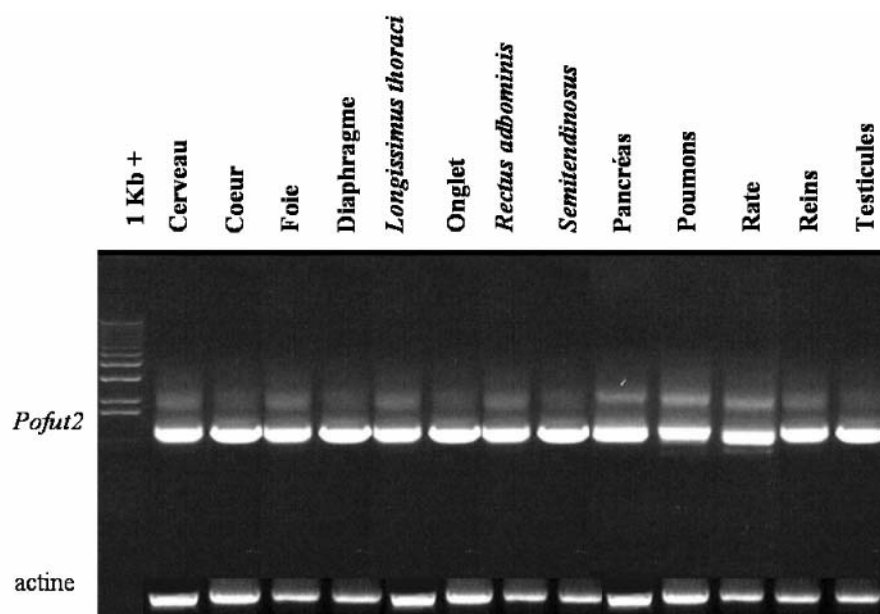


Figure 38 : Expression tissulaire des variants transcriptionnels de *Pofut2* dans les tissus de bovin adulte.

Les amorces utilisées sont référencées dans la table I de la publication 1. Des PCR réalisées avec un couple d'amorce spécifique de l'actine permettent de vérifier la qualité des ARNm ayant servi à la transcription inverse.

Chaque amplification obtenue est en fait un mélange de différents variants transcriptionnels du gène *Pofut2*. Nous avons donc récupéré ces amplifiats pour chacun des tissus et nous les avons ensuite clonés dans le plasmide TOPO2.1. Ce dernier est un plasmide de type vecteur T. La présence d'une thymine non appariée et liée à la topoisomérase à chaque extrémité du vecteur linéarisé par l'enzyme *EcoRV*, endonucléase générant des extrémités à bord franc, permet le clonage d'un amplifiat du fait de l'ajout d'une adénine aux extrémités 3' des fragments d'ADN néosynthétisés par l'ADN polymérase que nous utilisons. Les différents inserts ne pouvant être distingués ni par leur taille ni par des profils de restriction différents, nous avons procédé, après transformation bactérienne et mini préparation d'ADN plasmidique, au séquençage d'un grand nombre de clones. Ce travail a permis de dresser le tableau présenté dans la publication 1 (Figure 5B/*Pofut2*).

Une étude similaire a été conduite sur huit tissus provenant d’embryons bovins de race Charolaise âgés d’environ 2 mois. Ceux-ci nous ont été donnés par les services vétérinaires de l’abattoir de Limoges. Les huit mêmes tissus ont été prélevés sur trois embryons et des analyses indépendantes ont été conduites. Quel que soit l’embryon, les résultats sont identiques. Les ARNm ont été extraits puis rétro-transcrits de la même façon que dans le cas des tissus de bovins adultes. Les résultats obtenus pour les deux gènes sont résumés dans la figure 39.

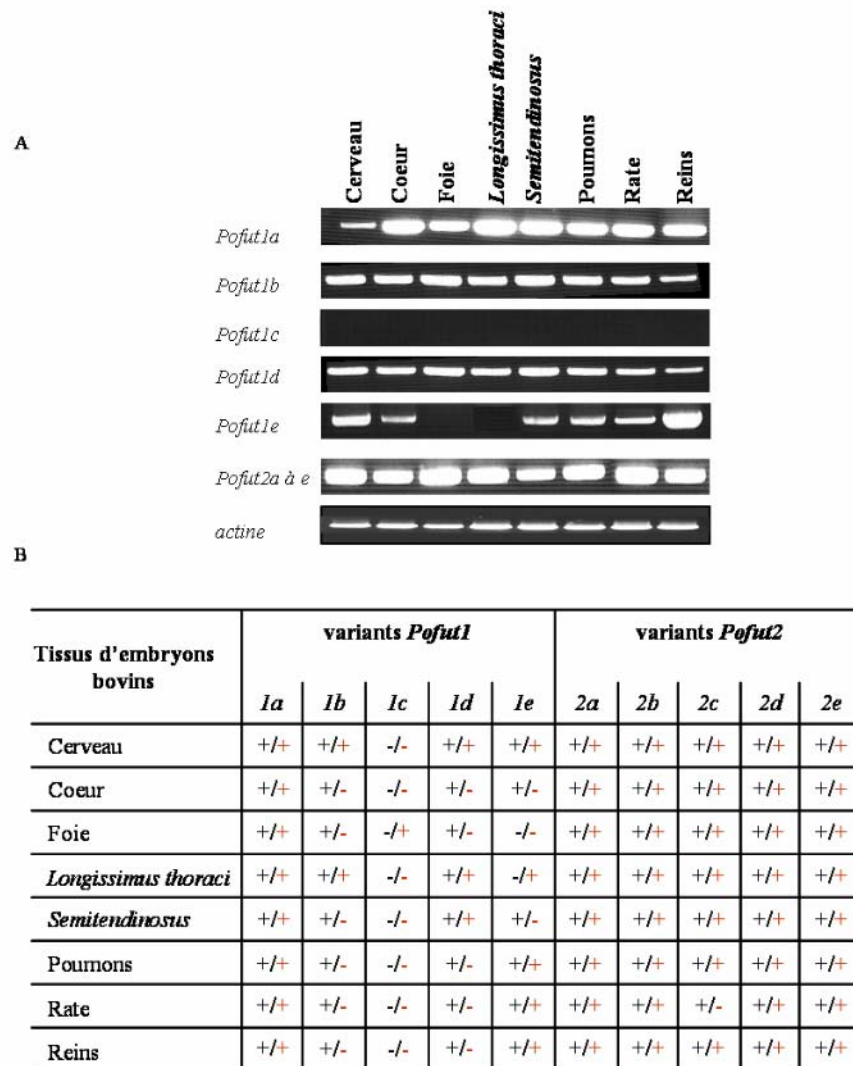


Figure 39 : Expression tissulaire des variants transcriptionnels de *Pofut1* et *Pofut2* dans les tissus d’embryon bovin. Comparaison avec l’expression tissulaire chez l’adulte.

Les amorces utilisées sont référencées dans la table I de la publication 1. A. Amplifications spécifiques des variants *a* à *e* de *Pofut1* et amplification commune des variants *a* à *e* de *Pofut2*. Les amplifications obtenues par PCR avec des amorces spécifiques de l’actine servent à contrôler la qualité des ARNm. B. Tableau récapitulatif l’expression des différents variants. Les signes en noir indiquent la présence (+) ou l’absence (-) des variants chez l’embryon; ces mêmes signes en rouge indiquent la présence ou l’absence des variants dans les mêmes tissus chez l’adulte.

La répartition des variants transcriptionnels dans les tissus d'embryons est différente de celle observée dans les tissus d'adultes :

> Comme pour les tissus adultes, le variant *Pofut1a* est présent dans tous les tissus embryonnaires testés.

> Le variant *Pofut1b* est présent dans tous les tissus embryonnaires testés alors que chez l'adulte, celui-ci n'est exprimé que dans le cerveau et le *Longissimus thoraci* (*Lt*, muscle squelettique du dos longeant la colonne vertébrale).

> Le variant *Pofut1c* qui était présent seulement dans le foie adulte, n'est détecté dans aucun des tissus d'embryon analysés.

> Le variant *Pofut1d* est présent dans les tissus embryonnaires étudiés, de façon ubiquitaire, alors que chez l'adulte, son expression est restreinte à 6 des 13 tissus : cerveau, diaphragme, *Lt*, ongle (muscle squelettique soutenant le diaphragme), *Semitendinosus* (muscle squelettique de l'arrière de la cuisse) et testicules.

> Le variant *Pofut1e* est absent du foie et du *Lt* embryonnaire. Chez l'adulte, il est aussi absent du foie mais également du *Semitendinosus* et du cœur.

> Tous les variants de *Pofut2* sont présents dans l'ensemble les tissus embryonnaires examinés. Chez l'adulte, ils ont aussi une expression ubiquiste, sauf dans le cas du variant *Pofut2b*, non détecté dans le *Rectus abdominis* (muscle squelettique de l'abdomen) et la rate.

Les profils d'expression tissulaire des transcrits sont donc différents selon le gène considéré, *Pofut1* ou *Pofut2*. De plus, le nombre de variants distincts de *Pofut* dénombrés par tissus est plus grand chez l'embryon par rapport à l'adulte. Ces résultats suggèrent une régulation transcriptionnelle différente, non seulement en fonction des tissus mais également du développement des animaux. Ces résultats sont en accord avec ceux obtenus chez la souris, du moins pour *Pofut1*. En effet, Shi et Stanley ont montré en 2003 par northern-blot, une augmentation du nombre et du niveau d'expression des transcrits du gène *Pofut1* dans des tissus embryonnaires par rapport à des tissus adultes. Tous ces résultats sont des arguments supplémentaires en faveur d'un rôle de la *O*-fucosylation durant l'embryogenèse.

Les variants transcriptionnels *Pofut1b* à *1e* et *Pofut2b* à *2e*, mis en évidence chez le bovin, semblent être une spécificité de cette espèce. En effet, des recherches dans les banques d'EST n'ont pas permis de trouver de séquences équivalentes chez d'autres organismes. Seuls deux EST humains, BP318733 et DB158060, obtenus respectivement à partir de tissus péricardique et de thymus, coderaient une protéine de 186 acides aminés équivalente aux protéines bovines *Pofut2C*, *D* et *E*. Ces protéines résultent d'un épissage alternatif qui élimine les 31 premières bases de l'exon 4 humain, et non les 61 premières comme pour les variants *Pofut2c*, *d* et *e* bovin. Chez l'Homme, ce phénomène conduit à l'utilisation d'un codon stop prématuré à la fin de ce même exon (Fig.40.) et non dans l'exon 5 comme dans le cas du gène bovin (cf. Publication 1, figure 4B). Ces deux EST étant seulement séquencés dans leurs parties 5', il est impossible de se prononcer sur les épissages se produisant après l'exon 4 humain.

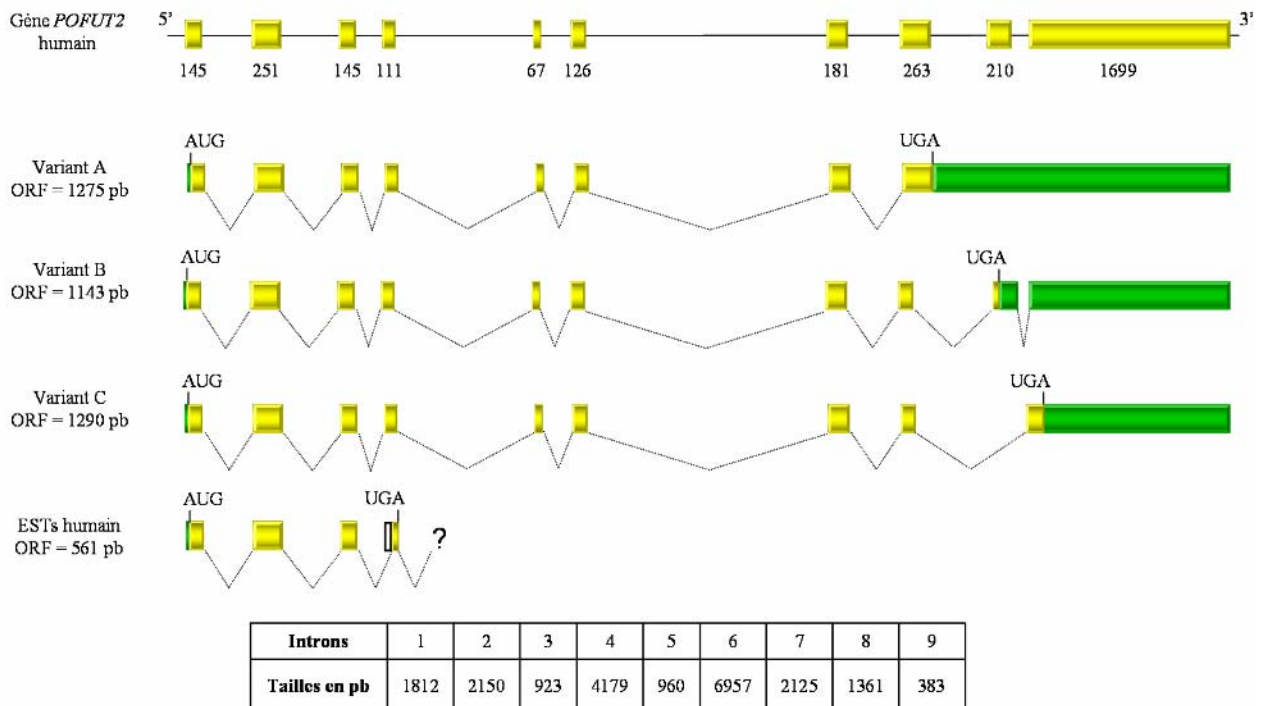


Figure 40 : Représentation schématique de l'organisation génomique du gène *POFUT2* humain et de l'épissage alternatif conduisant aux transcrits référencés dans les bases de données moléculaires (variants A, B et C, et ESTs).

Les régions non traduites des transcrits sont en vert, les tailles des exons sont indiquées sous la structure du gène (en pb), les tailles des introns sont notées dans le tableau ci-dessus.

1.2.2 Production d'anticorps anti-Pofut bovins

Avant de déterminer les activités des *O*-fucosyltransférases bovines codées par les différents variants, et dans le but d'étudier l'expression de celles-ci dans différents tissus bovins ou autres, nous avons produit des anticorps spécifiquement dirigés contre ces enzymes.

Les séquences codantes des transcrits *Pofut1a* et *Pofut2b* (uniques variants dont nous disposions à l'époque) ont été clonées de manière orientée dans le vecteur d'expression bactérien pET28c, grâce aux sites de restriction *NdeI* et *XhoI*. Le clonage des séquences bovines a nécessité, au préalable, leur amplification avec des couples d'amorces apportant les sites de restriction enzymatique et permettant de conserver le cadre de lecture avec le vecteur pET28c qui apporte le codon d'initiation de la traduction. Les couples d'amorces utilisés sont répertoriés dans le tableau 3.

Tableau 3 : Couples d'amorces utilisés pour les clonages en vecteur d'expression procaryote pET28c. Les sites de restrictions sont soulignés.

Amorces	Séquences
Pofut1- <i>NdeI</i> /5'	5'- <u>CATATGGGCGCCGCGCG</u> -3'
Pofut1- <i>XhoI</i> /3'	5'- <u>CTCGAGGAATTCATCCCGCAGCTG</u> -3'
Pofut2- <i>NdeI</i> /5'	5'- <u>CATATGGCGGGCGCTTGCCGTCGTC</u> -3'
Pofut2- <i>XhoI</i> /3'	5'- <u>CTCGAGGTATGCAATCCTCCAATG</u> -3'

Les amplifiats ainsi obtenus sont d'abord clonés dans le plasmide TOPO2.1. Puis, les séquences d'intérêt sont isolées grâce à la double digestion du plasmide par les enzymes de restriction *NdeI* et *XhoI*. Elles sont ensuite introduites aux mêmes sites dans le vecteur pET28c, préalablement digéré. Ce vecteur porte le promoteur du bactériophage T7 et l'opérateur lacUV5 pour le contrôle de l'expression du gène d'intérêt. Les constructions, pET28c-*Pofut1a* et pET28c-*Pofut2b*, sont transformées dans la souche *Escherichia coli* BL21(DE3) qui porte dans son génome le gène codant l'ARN polymérase T7 (lysogène λ DE3) sous la dépendance du promoteur lacUV5, inductible par un galactoside. En présence de l'inducteur, un β galactoside, le gène de l'ARN polymérase présent sur le chromosome bactérien s'exprime. L'absence de répression permet à cette polymérase d'initier la transcription du gène d'intérêt.

L'expression de chaque protéine est induite pendant 4 heures à 37°C par l'addition d'IPTG à une concentration de 1mM. L'analyse, par électrophorèse en gel d'acrylamide à 10,5% en conditions dénaturantes, des fractions protéiques totales avant (Fig.41., pistes 1) et après induction (Fig.41., Pistes 2) permet de mettre en évidence l'expression des deux enzymes. Celles-ci sont ensuite purifiées (Fig.41., Pistes 3) sur résine de Nickel grâce à la présence de six résidus histidine à chacune de leurs extrémités (Fig.42.). Les tailles attendues pour les protéines taggées et non glycosylées, car exprimées dans la bactérie, sont pour Pofut1A de 46,9 KDa et pour Pofut2B de 52 KDa.

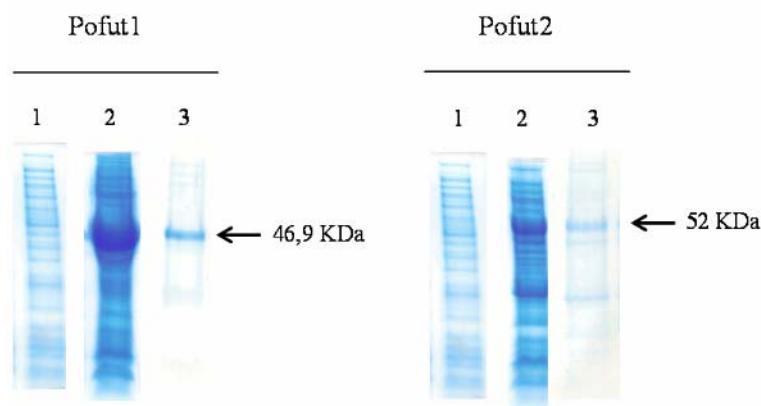


Figure 41 : Analyse par SDS-PAGE 10,5% de l'expression des enzymes recombinantes bovines Pofut1A et Pofut2B.

Pistes 1, protéines bactériennes avant induction; pistes 2, protéines bactériennes après induction; pistes 3, 1µg des fractions protéiques purifiées.

La présence de deux étiquettes histidine, à la fois en N et C-terminal, sur chaque protéine permet d'augmenter le rendement des étapes de purification sur résine de nickel (Ni^{2+}).

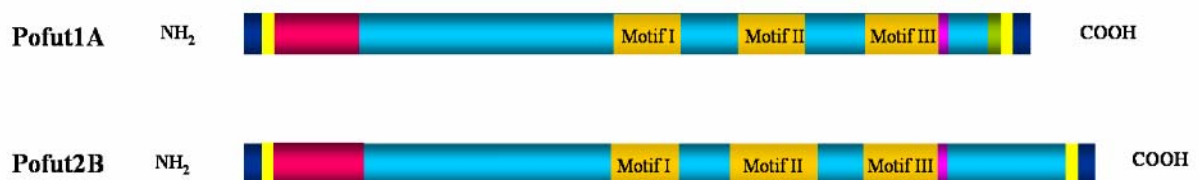


Figure 42 : Représentation schématique des protéines Pofut1A et Pofut2B bovines produites dans le système bactérien.

Le peptide signal est en rose, les motifs conservés entre les α 2-, α 6- et les *O*-fucosyltransférases en orange, le DXD en violet, le KDEL-like de Pofut1 en vert. Les acides aminés, au nombre de 51, rajoutés par le vecteur sont en jaune et les étiquettes histidine sont en bleu foncé.

Après purification de quantités avoisinant 1mg de protéines recombinantes, Pofut1 et Pofut2 bovines ont été utilisées pour immuniser des lapins (2 individus pour chaque protéine), en vue d'obtenir des anticorps polyclonaux. Le programme d'immunisation a été réalisé par une société spécialisée (Agro-Bio, La Ferté St Aubin, France).

Pour tester la reconnaissance spécifique de ces anticorps, nous avons effectué des analyses par western blot avec chaque serum reçu, d'une part sur les protéines ayant servi à l'immunisation des lapins, et d'autre part sur des lysats de cellules COS-1 transfectées avec le vecteur d'expression eucaryote pcDNA/TOPO3.1 contenant les inserts *Pofut1a* ou *Pofut2a* ou *Pofut2b*. Comme c'est le variant *Pofut2a* qui code potentiellement la forme active de l'enzyme, il est impératif de l'inclure dans nos tests. La figure 43 présente les résultats obtenus pour chaque protéine avec le dernier sérum d'un des deux lapins immunisés. Le contrôle par western blot des sérums pré-immuns des lapins est négatif (résultats non présentés).

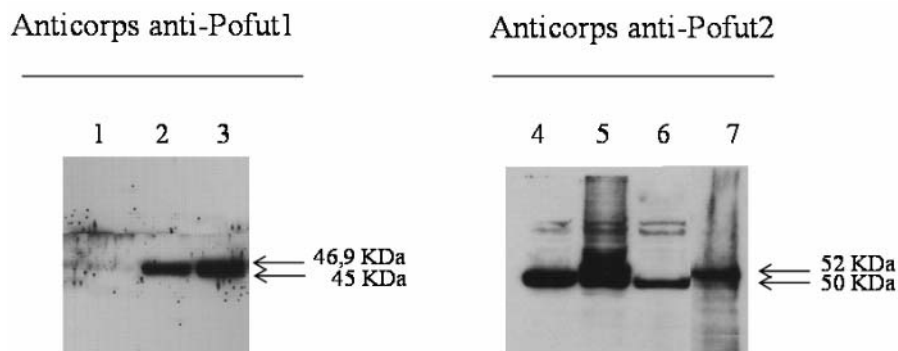


Figure 43 : Vérification de la spécificité des anticorps dirigés contre les enzymes bovines Pofut1 et Pofut2.

Pistes 1 et 6, 50µg de protéines de cellules COS-1 non transfectées; piste 2, 50µg de protéines de cellules COS-1 transfectées avec le vecteur pcDNA/TOPO3.1-*Pofut1a*; piste 3, 100ng de protéine Pofut1A exprimée dans *E. coli* et purifiée; piste 4, 50µg de protéines de cellules COS-1 transfectées avec le vecteur pcDNA/TOPO3.1-*Pofut2a*; piste 5, 50µg de protéines de cellules COS-1 transfectées avec le vecteur pcDNA/TOPO3.1-*Pofut2b*; piste 7, 100ng de protéine Pofut2B exprimée dans *E. coli* et purifiée.

Les protéines produites dans les cellules COS-1 (Fig.44.) sont reconnues par les anticorps correspondants. Ces enzymes potentiellement glycosylées (comme nous l'avons vu dans le chapitre 1) et dépourvues de leur peptide signal (masse d'environ 3 KDa) migrent au même niveau que la protéine produite en système bactérien, qui conserve le peptide signal, est dépourvue de glycosylation et possède deux étiquettes histidine.

La masse moléculaire attendue, déduite de la séquence peptidique après clivage du peptide signal, est pour Pofut1A de 41 KDa, à laquelle il faut ajouter la masse moléculaire des *N*-glycanes potentiellement présents sur cette enzyme réticulaire, soit 4 KDa (Wang et Spellman, 1998). Les masses moléculaires attendues, déduites des séquences peptidiques après clivage du peptide signal sont pour Pofut2A de 44 KDa et pour Pofut2B de 46 KDa. La protéine Pofut2A bovine, comme ses orthologues Mammaliennes, possède trois sites potentiels de *N*-glycosylation. Aucune étude n'a, pour l'instant, porté sur le nombre et la composition des *N*-glycanes présents sur l'enzyme Pofut2. Néanmoins, des hypothèses ont été émises sur sa localisation réticulaire chez l'Homme et la souris (Luo et Haltiwanger, 2005). Il est donc probable que, si celle-ci est glycosylée, ses *N*-glycanes soient également de type oligomannosidique voire hybride. Si on admet que la masse moléculaire de chaque *N*-glycane est de 2 KDa (masse moléculaire approximative de chaque *N*-glycane de Pofut1), la protéine recombinante de 50 KDa révélée dans les extraits de cellules COS-1 transfectées avec la construction pcDNA/TOPO3.1-*Pofut2a* est une forme triglycosylée de l'enzyme. La protéine recombinante de 52 KDa révélée dans les extraits de cellules transfectées avec la construction pcDNA/TOPO3.1-*Pofut2b* est une forme triglycosylée de l'enzyme Pofut2B qui possède également les 3 sites potentiels de *N*-glycosylation.

Etant donné les fortes homologies entre les *O*-fucosyltransférases de différentes espèces de Mammifères, il n'est pas étonnant que nos anticorps croisent avec les enzymes endogènes présentes dans les cellules COS-1 (cellules de rein de singe vert). Toutefois, dans ces cellules, l'enzyme Pofut1 semble plus faiblement reconnue que l'enzyme Pofut2. Seule une très fine bande est visible. Toutefois, si une quantité deux fois supérieure d'extrait de protéines issues des cellules COS-1 est déposée, une bande de plus forte intensité est alors marquée (voir figure 47B, piste 1).

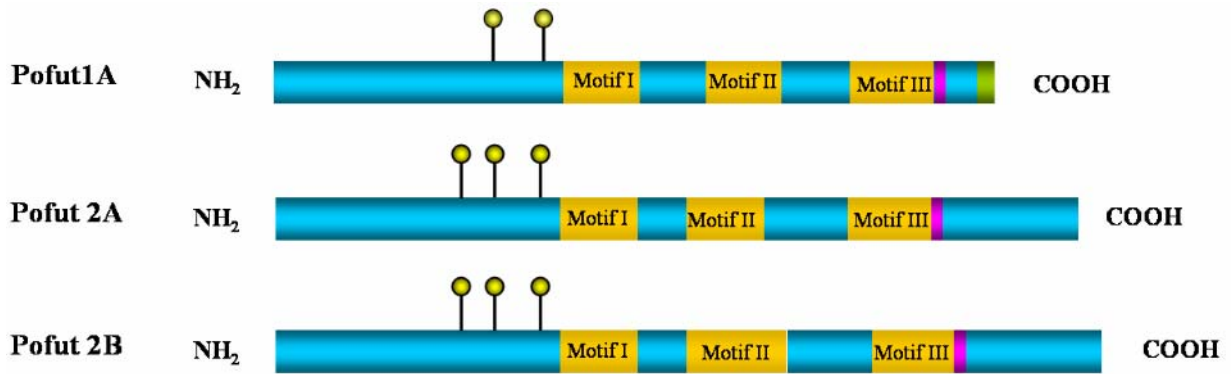


Figure 44 : Représentation schématique des protéines bovines Pofut1A, Pofut2A et Pofut2B produites dans les cellules COS-1.

Les motifs conservés entre les $\alpha 2$ -, $\alpha 6$ - et les *O*-fucosyltransférases sont en orange, le DXD en violet, le KDEL-like de Pofut1 en vert, les sites potentiels de *N*-glycosylation en jaune. Le motif II de l'enzyme Pofut2B présente 30 acides aminés supplémentaires par rapport à Pofut2A.

1.2.3 Activités des protéines codées par les différents variants

Parmi les différents variants transcriptionnels mis en évidence pour les deux gènes bovins *Pofut1* et *Pofut2*, seuls les variants *Pofut1a* et *Pofut2a* codent respectivement les protéines homologues à celles caractérisées chez l'Homme et la souris. Les autres variants coderaient en effet des protéines plus ou moins tronquées (sauf *Pofut2b* qui code une protéine plus grande que *Pofut2a*), dépourvues de la plupart des motifs conservés et donc potentiellement inactives. Les figures 45 et 46 présentent les alignements de séquences des protéines codées par les variants *Pofut1a* à *Pofut1e* et *Pofut2a* à *Pofut2e*.

Pofut1A	MGAAAWAPSL	LPPRVSLLLL	LLPLPGLPVG	SWDPAGYLLY	CPCMGRFGNQ	ADHFLGSLAF	60
Pofut1B	MGAAAWAPSL	LPPRVSLLLL	LLPLPGLPVG	SWDPAGYLLY	CPCMGRFGNQ	ADHFLGSLAF	60
Pofut1C	MGAAAWAPSL	LPPRVSLLLL	LLPLPGLPVG	SWDPAGYLLY	CPCMGRFGNQ	ADHFLGSLAF	60
Pofut1D	MGAAAWAPSL	LPPRVSLLLL	LLPLPGLPVG	SWDPAGYLLY	CPCMGRFGNQ	ADHFLGSLAF	60
Pofut1E	MGAAAWAPSL	LPPRVSLLLL	LLPLPGLPVG	SWDPAGYLLY	CPCMGRFGNQ	ADHFLGSLAF	60
Pofut1A	AKLLNRTLAV	PPWIEYQHKK	PPFTNVHVSYS	QKYFKLEPLQ	AYHRVISLED	FMEKLAPTHW	120
Pofut1B	AKLLNRTLAV	PPWIEYQHKK	PPFTNVHVSYS	QKYFKLEPLQ	AYHRVISLED	FMEKLAPTHW	120
Pofut1C	AKLLNRTLAV	PPWIEYQHHE	PPFTNVHVSYS	QKYFKLEPLQ	AYHRVISLED	FMEKLAPTHW	120
Pofut1D	AKLLNRTLAV	PPWIEYQHKK	PPFTNVHVSYS	QKYFKLEPLQ	AYHRVISLED	FMEKLAPTHW	120
Pofut1E	AKLLNRTLAV	PPWIEYQHKK	PPFTNVHVSYS	QKYFKLEPLQ	AYHRVISLED	FMEKLAPTHW	120
Pofut1A	PPEKRVAYCF	EVAQRSPDK	KTCMPKEGNP	FGPFWDQFHV	SFNKSELFAG	ISFSASYKDQ	180
Pofut1B	PPEKRVAYCF	EVAQRSPDK	KTCMPKEGNP	FGPFWDQFHV	SFNKSELFAG	ISFSASYKDQ	180
Pofut1C	PPEKRVAYCS	EVAARRSPDK	KTCMPKEGNP	FGPFWDQFHV	SFNKSELFAG	ISFSASYKDQ	180
Pofut1D	PPEKRVAYCF	EVAQRSPDK	KTCMPKEGNP	FGPFWDQFHV	SFNKSELFAG	ISFSASYKDQ	180
Pofut1E	PPEKRVAYCF	EVAQRSPDK	KTCMPKEGNP	FGPFWDQFHV	SFNKSELFAG	ISFSASYKDQ	180
Pofut1A	WIQRFSPPEH	PVLALPGAPA	QFPVLEEHRP	LQKYMVWVSD	MVRTGEAQIH	AHLIRPYVGI	240
Pofut1B	WIQRFSPPEH	PVLALPGAPA	QFPVLEEHRP	LQKYMVWVSD	MVRTGEAQIH	AHLIRPYVGI	240
Pofut1C	WIQR TLVPHF	HSKHQSPHWA	QRSASVWLLT	ACHVSEDVAS	RSNQDGCVS	-----	229
Pofut1D	WIQR-----	-----	-----	-----	-----	-----	184
Pofut1E	WIQR A -----	-----	-----	-----	-----	-----	185
Pofut1A	HLRIGSDWKN	ACAMLKDGTA	GAHFMASPQC	VGYSRHTTAP	LTMTMCLPDL	KEIRRALKLW	300
Pofut1B	HLRIGSDWKN	ACAMLKDGTA	GAHFMASPQC	VGYSRHTTAP	LTMTMCLPDL	KEIRRALKLW	300
Pofut1C	-----	-----	-----	-----	-----	-----	-----
Pofut1D	-----	-----	-----	-----	-----	-----	-----
Pofut1E	-----	-----	-----	-----	-----	-----	-----
Pofut1A	VTALNAQSVY	IATDSESYLP	EIQQLFKGKV	KVVS L KPEVA	Q IDL Y ILGQA	D HFIGNCVSS	360
Pofut1B	VTALNAQSVY	IATDSESYLP	EIQQLFKGKS	L SFAV P KASE	M SLL L GDRKG	K -----	351
Pofut1C	-----	-----	-----	-----	-----	-----	-----
Pofut1D	-----	-----	-----	-----	-----	-----	-----
Pofut1E	-----	-----	-----	-----	-----	-----	-----
Pofut1A	F TAFV K RE D	LHGRPSSFFG	MDRPPQL R DE F	391			
Pofut1B	-----	-----	-----				
Pofut1C	-----	-----	-----				
Pofut1D	-----	-----	-----				
Pofut1E	-----	-----	-----				

Figure 45 : Alignement des séquences protéiques déduites des cinq variants transcriptionnels du gène bovin *Pofut1*.

Le peptide signal est en rose, les sites potentiels de *N*-glycosylation en jaune, les trois motifs conservés avec les α 2- et α 6-fucosyltransférases sont : en bleu (motif I), en gris (motif II) et en vert (motif III). Le motif DXD est en gras souligné et le domaine de rétention dans le RE de type KDEL-like est en vert foncé. La position des acides aminés est indiquée sur la droite de l'alignement. Les acides aminés qui diffèrent de ceux de Pofut1A sont en rouge.

Pofut2A	MAALAVVFL	LWAASWPSAS	AFGEEFWPGQ	SAADILSGAA	SRRRYLLYDV	NPPEGFNLR	60
Pofut2B	MAALAVVFL	LWAASWPSAS	AFGEEFWPGQ	SAADILSGAA	SRRRYLLYDV	NPPEGFNLR	60
Pofut2C	MAALAVVFL	LWAASWPSAS	AFGEEFWPGQ	SAADILSGAA	SRRRYLLYDV	NPPEGFNLR	60
Pofut2D	MAALAVVFL	LWAASWPSAS	AFGEEFWPGQ	SAADILSGAA	SRRRYLLYDV	NPPEGFNLR	60
Pofut2E	MAALAVVFL	LWAASWPSAS	AFGEEFWPGQ	SAADILSGAA	SRRRYLLYDV	NPPEGFNLR	60
Pofut2A	DVYIRIASLL	KTLLKTEEWV	LVLPPWGRLY	HWQSPDIHQV	RIPWSDFFDL	PSLNRNIPVI	120
Pofut2B	DVYIRIASLL	KTLLKTEEWV	LVLPPWGRLY	HWQSPDIHQV	RIPWSDFFDL	PSLNRNIPVI	120
Pofut2C	DVYIRIASLL	KTLLKTEEWV	LVLPPWGRLY	HWQSPDIHQV	RIPWSDFFDL	PSLNRNIPVI	120
Pofut2D	DVYIRIASLL	KTLLKTEEWV	LVLPPWGRLY	HWQSPDIHQV	RIPWSDFFDL	PSLNRNIPVI	120
Pofut2E	DVYIRIASLL	KTLLKTEEWV	LVLPPWGRLY	HWQSPDIHQV	RIPWSDFFDL	PSLNRNIPVI	120
Pofut2A	EYEQFIAESG	GPFIDQVYIL	QGYAEGWKEG	AWEEKIDSRP	CIEPPLYSQD	KHEYRGWFW	180
Pofut2B	EYEQFIAESG	GPFIDQVYIL	QGYAEGWKEG	AWEEKIDSRP	CIEPPLYSQD	KHEYRGWFW	180
Pofut2C	EYEQFIAESG	GPFIDQVYIL	QGYAEGWKEG	AWEEKIDSRP	CIEPPLYSQD	KHEYRDRPP	180
Pofut2D	EYEQFIAESG	GPFIDQVYIL	QGYAEGWKEG	AWEEKIDSRP	CIEPPLYSQD	KHEYRDRPP	180
Pofut2E	EYEQFIAESG	GPFIDQVYIL	QGYAEGWKEG	AWEEKIDSRP	CIEPPLYSQD	KHEYRDRPP	180
Pofut2A	GYEETRGLNV	SCLSVQGSAS	IIAPVLLRNT	SARVMDRA	ENLLHDHYGG	KEYWDTRRSM	240
Pofut2B	GYEETRGLNV	SCLSVQGSAS	IIAPVLLRNT	SARVMDRA	ENLLHDHYGG	KEYWDTRRSM	240
Pofut2C	SSRLYY----	-----	-----	-----	-----	-----	186
Pofut2D	SSRLYY----	-----	-----	-----	-----	-----	186
Pofut2E	SSRLYY----	-----	-----	-----	-----	-----	186
Pofut2A	VFARHLRAVG	DEFRSRYLNS	TDVADGIPFE	EDWTQMKVKL	GSSLGGPYLG	VHLRRKDFIW	300
Pofut2B	VFARHLRAVG	DEFRSRYLNS	TDVADGIPFE	EDWTQMKVKL	GSSLGGPYLG	VHLRRKDFIW	300
Pofut2C	-----	-----	-----	-----	-----	-----	
Pofut2D	-----	-----	-----	-----	-----	-----	
Pofut2E	-----	-----	-----	-----	-----	-----	
Pofut2A	GHREDVPSLD	GAVRKIRSLM	KAHQLDKVFV	ATDAVRK----	-----	-----	337
Pofut2B	GHREDVPSLD	GAVRKIRSLM	KAHQLDKVFV	ATDAVRKELW	VCACLCETPL	GCASPPRCPL	360
Pofut2C	-----	-----	-----	-----	-----	-----	
Pofut2D	-----	-----	-----	-----	-----	-----	
Pofut2E	-----	-----	-----	-----	-----	-----	
Pofut2A	-----ELE	ELRLLPEMV	RFEPTWEELE	LYKDGGVAIV	DQWICAHARE	FIGTSVSTFS	390
Pofut2B	TSPLAPTELE	ELRLLPEMV	RFEPTWEELE	LYKDGGVAIV	DQWICAHARE	FIGTSVSTFS	420
Pofut2C	-----	-----	-----	-----	-----	-----	
Pofut2D	-----	-----	-----	-----	-----	-----	
Pofut2E	-----	-----	-----	-----	-----	-----	
Pofut2A	FRIHEERE IL	GLDPKTTYNR	FCGDKEKACE	QPTHWRIAY	429		
Pofut2B	FRIHEERE IL	GLDPKTTYNR	FCGDKEKACE	QPTHWRIAY	459		
Pofut2C	-----	-----	-----	-----			
Pofut2D	-----	-----	-----	-----			
Pofut2E	-----	-----	-----	-----			

Figure 46 : Alignement des séquences protéiques déduites des cinq variants transcriptionnels du gène bovin *Pofut2*.

Le peptide signal est en rose, les sites potentiels de *N*-glycosylation en jaune, les trois motifs conservés avec les α 2- et α 6-fucosyltransférases sont : en bleu (motif I), en gris (motif II) et en vert (motif III). Le motif DXD est en gras souligné. Dans la protéine Pofut2B, l'insertion de 30aa dans le motif II est indiquée en gras. La position des acides aminés est indiquée sur la droite de l'alignement. Les acides aminés qui diffèrent de ceux de Pofut2A sont en rouge.

Dans le but de déterminer si les variants transcriptionnels sont susceptibles de coder des protéines dotées d'une activité *O*-fucosyltransférase, des tests d'activité *in vitro* ont été effectués. La séquence codante de chacun des variants transcriptionnels a été clonée dans le vecteur d'expression eucaryote pcDNA/TOPO3.1. Les plasmides recombinants ont ensuite été utilisés pour transférer des cellules COS-1. Les tests d'activité sont réalisés avec 5 μ g de lysats cellulaires en utilisant comme substrat donneur le GDP-fucose radiomarqué au C¹⁴ et comme substrats accepteurs, le domaine EGF du facteur de coagulation IX humain pour Pofut1 (Fig.47.) ou le troisième domaine TSR de la thrombospondine I humaine pour Pofut2 (Fig.48.).

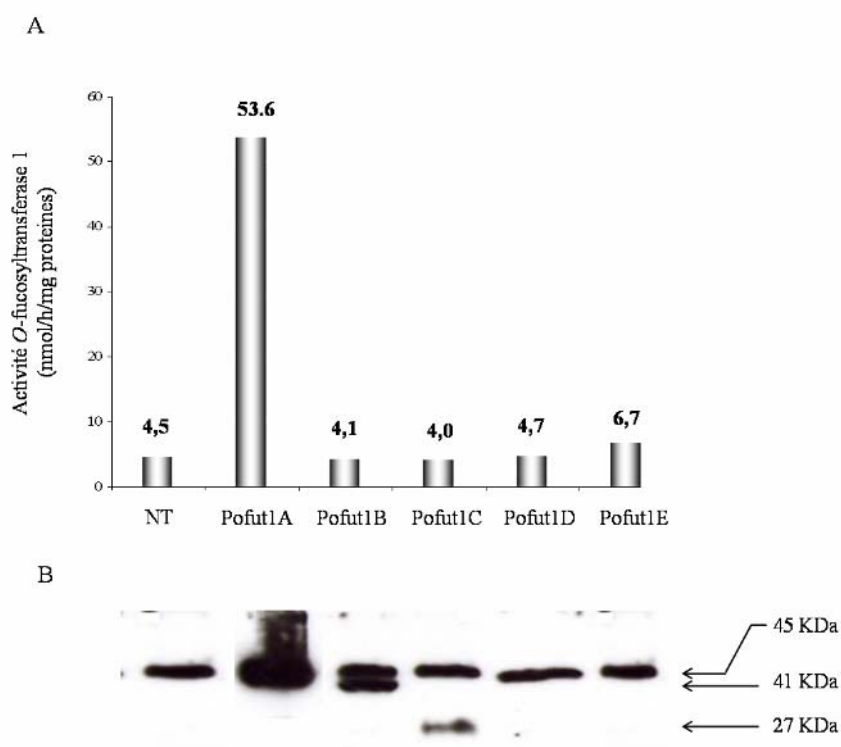


Figure 47 : Activité *O*-fucosyltransférase 1 des protéines codées par les variants bovins *Pofut1a* à *Pofut1e*.

A, Les activités sont données en nmole par heure et par mg de protéines totales. Elles sont indiquées au dessus de chaque histogramme; B, Analyses en western blot de l'expression des protéines recombinantes dans les cellules COS-1 et détection à l'aide de l'anticorps anti-Pofut1. 100 μ g de protéines totales ont été déposées; les tailles des protéines détectées sont indiquées.

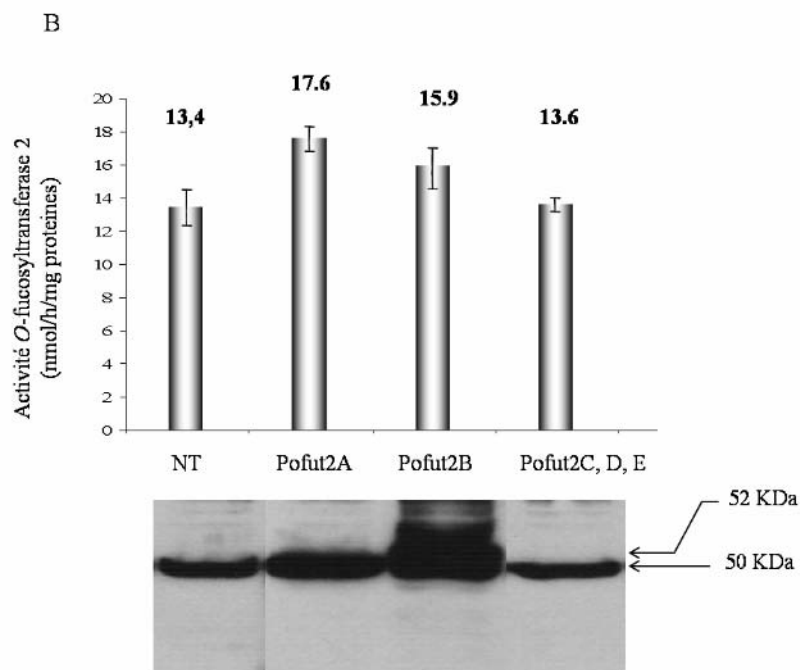


Figure 48 : Activité *O*-fucosyltransférase 2 des protéines codées par les variants bovins *Pofut2a* à *Pofut2e*.

A, Les activités sont données en nmole par heure et par mg de protéines totales. Elles sont indiquées au dessus de chaque histogramme. B, Analyses en western blot de l'expression des protéines recombinantes dans les cellules COS-1 et détection à l'aide de l'anticorps anti-Pofut2. 50µg de protéines totales ont été déposées; les tailles des protéines détectées sont indiquées.

Dans le cas de *Pofut1*, seul le variant *Pofut1a* code une protéine *O*-fucosyltransférase 1 active. Dans le cas de *Pofut2*, seul le variant *Pofut2a* code une protéine ayant une activité significative par rapport à l'activité endogène détectée dans les cellules COS-1. Le variant *Pofut2b* semble coder une enzyme ayant une activité *O*-fucosyltransférase 2. Néanmoins, l'activité de celle-ci, dans les conditions du test *in vitro*, n'est pas significative. Les autres variants de *Pofut2* codent tous la même protéine tronquée dépourvue d'activité *O*-fucosyltransférase.

Afin de vérifier l'expression des protéines recombinantes, nous avons effectué des analyses par western blot, avec nos anticorps polyclonaux, sur les lysats protéiques des cellules COS-1 transfectées. Dans la figure 47, l'anticorps anti-Pofut1A bovin reconnaît les protéines Pofut1A, B et C à leurs masses moléculaires prédites, respectivement 45, 41 et 27 KDa (sans le peptide signal et diglycosylées). Cependant, l'anticorps ne reconnaît pas les protéines D et E, dont la masse moléculaire prédite (sans le peptide signal et diglycosylées) est de 22 KDa.

Dans la figure 48, l'anticorps anti-Pofut2B bovin reconnaît les protéines Pofut2A et B à leurs masses moléculaires prédites, respectivement 50 et 52 KDa (sans le peptide signal et triglycosylées). Cependant, l'anticorps ne reconnaît pas les protéines C, D et E, qui sont identiques et d'une masse moléculaire prédite (sans le peptide signal et non glycosylée, ne possédant aucun des 3 sites potentiels de *N*-glycosylation) est de 18 KDa.

Les expériences de transfection ayant été conduites de façon simultanée, on peut penser que les protéines Pofut1D et E ainsi que les protéines Pofut2C, D et E, bien qu'étant dépourvues de la majeure partie de leur domaine catalytique, ont été correctement produites, mais sont trop tronquées pour être reconnues par les anticorps.

1.2.4 Evolution des gènes de *O*-fucosyltransférases

Une recherche *in silico* dans les bases de données moléculaires a permis la sélection de 27 séquences protéiques complètes correspondants à Pofut1 et 21 pour Pofut2, provenant d'espèces appartenant à divers groupes taxonomiques représentatifs du monde Animal. Dans chacun des génomes étudiés, une seule copie de chacun des gènes *Pofut1* et *Pofut2* est présente. Les séquences peptidiques déduites ont été alignées en utilisant comme références les séquences des protéines bovines actives Pofut1A et Pofut2A. Les annexes I et II à la fin de ce chapitre sont dédiées à la présentation de ces alignements.

Les tailles des séquences utilisées vont de 380 à 404 acides aminés pour Pofut1 et de 405 à 495 acides aminés pour Pofut2. Les différences de tailles pour chacune des enzymes sont surtout dues à des divergences au niveau de la partie N-terminale des protéines, incluant le peptide signal. Ainsi, la partie N-terminale des séquences, les portions de l'alignement comportant des insertions-délétions (indels) chez certaines espèces et celles dont l'alignement peut être discutable n'ont pas été prises en compte pour la construction des arbres phylogénétiques.

Dans toutes les séquences, les motifs I, II et III, communs avec les $\alpha 2$ et $\alpha 6$ -fucosyltransférases, sont conservés et des sites de *N*-glycosylation sont retrouvés. Les enzymes Pofut1 et Pofut2 sont donc toutes potentiellement des glycoprotéines. Les séquences correspondant à Pofut1 possèdent toutes en C-terminal une séquence de type KDEL-like, sauf celle de l'Echinoderme *Strongylocentrotus purpuratus* qui est SIGV. Les sites de *N*-glycosylation de Pofut1 sont au nombre de deux chez les Mammifères et leurs positions sont strictement conservées. A l'échelle du monde Animal, seul le premier site est conservé sauf

pour les nématodes *Caenorhabditis sp.* et *Strongyloides ratti*, pour l’Insecte *Homalodisca coagulata* et pour le cnidaire *Acropora millepora* (pour ces trois dernières espèces, les séquences de Pofut1 sont partielles). La portion d’alignement des séquences de Pofut1 chez les Animaux où sont situés les sites de *N*-glycosylation conservés chez les Mammifères, est présentée dans la publication 2. En ce qui concerne Pofut2, les deux premiers sites potentiels de *N*-glycosylation sont retrouvés à des positions conservées chez les Ostéichthyens. Le site 2 est aussi présent chez les Urochordés (*Ciona*) et chez *Caenorhabditis sp.*. Le site 3 semble moins conservé puisque à l’heure actuelle, il n’est présent que chez les Mammifères, les Téléostéens et l’Insecte, *Anopheles gambiae*.

Grâce au logiciel smart (<http://smart.embl-heidelberg.de/>), nous avons identifié les peptides signaux potentiels des protéines Pofut1 et Pofut2 de toutes les espèces choisies (Voir annexes I et II). Le fait que ces peptides soient présents à l’extrémité N-terminale de pratiquement toutes les séquences analysées (sauf pour Pofut1 et Pofut2 du poisson zèbre *Danio rerio*) ainsi que la démonstration faite de la localisation des enzymes Pofut1 et Pofut2 humaines (Luo et Haltiwanger, 2005; Luo, Nita-Lazar *et al.*, 2006), suggèrent une localisation réticulaire des *O*-fucosyltransférases chez toutes les espèces. Ceci est appuyé par les résultats obtenus par le logiciel Predotar (<http://urgi.infobiogen.fr/predotar/french.html>) qui prédit une localisation réticulaire pour l’ensemble des protéines étudiées ici.

L’analyse et la comparaison de toutes ces séquences nous ont amené à établir les organisations génomiques de *Pofut1* et *Pofut2* dans la majorité des espèces étudiées. Certaines de ces organisations sont déjà présentées dans la figure 7 de la publication 1. Les figures 49 et 50 incluent, en plus des organisations génomiques présentées dans l’article, celles des gènes *Pofut1* et *Pofut2* d’autres organismes pour lesquels nous possédons des informations sur les CDS et sur les séquences génomiques. Ces représentations tiennent compte uniquement des exons codant la forme potentiellement active des enzymes; seuls les CDS sont indiqués.

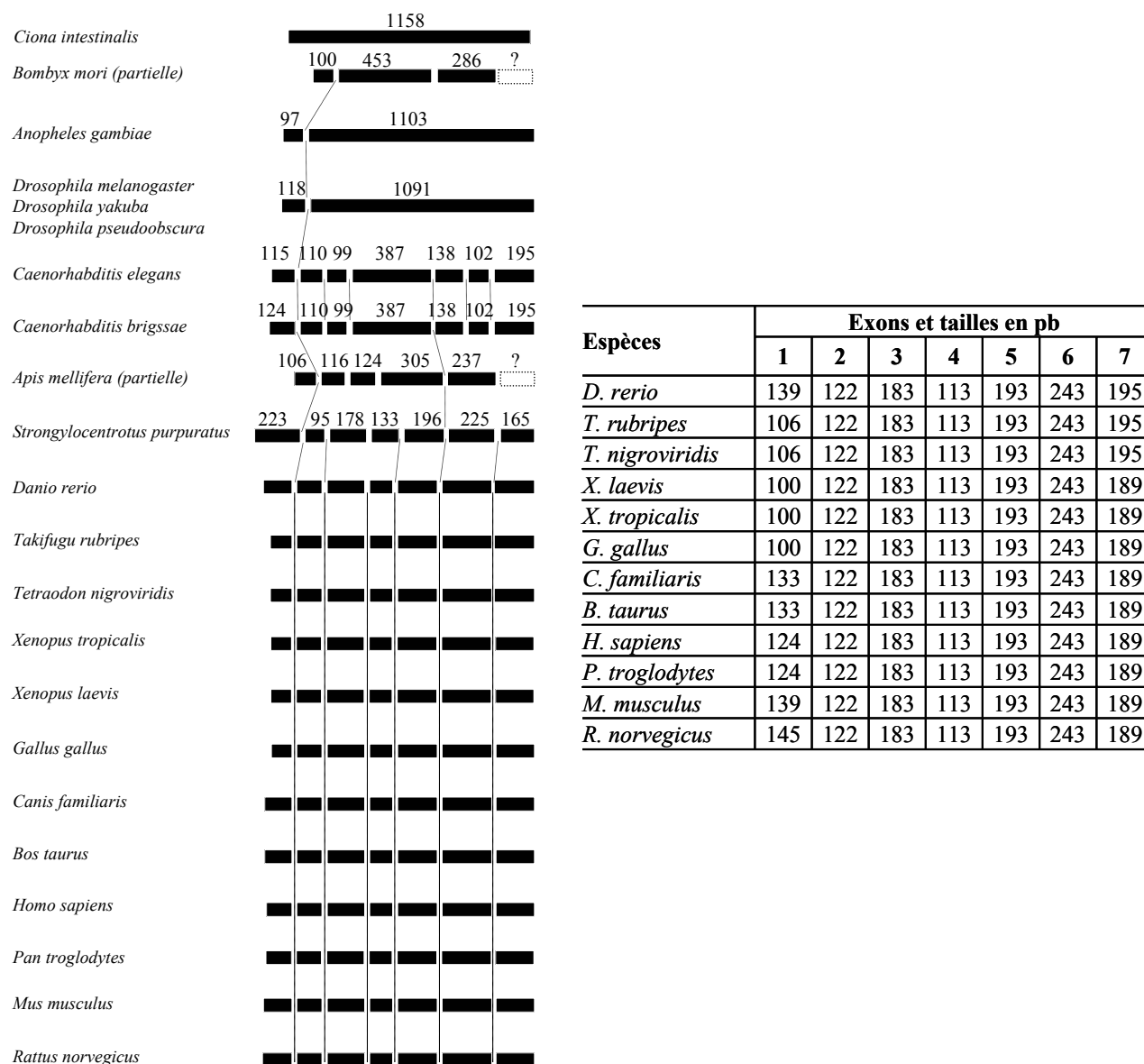


Figure 49 : Organisations exon/intron des CDS des gènes *Pofut1* chez différentes espèces animales.

Pour les Vertébrés, les tailles des exons homologues sont indiquées dans le tableau. Celles des invertébrés sont indiquées au dessus de chaque structure.

- Pour *Pofut1*, les organisations génomiques en sept exons codants sont retrouvées chez tous les Vertébrés. La position des six introns est strictement conservée.

> Six introns sont également présents chez l’Echinoderme *Strongylocentrotus purpuratus*. Néanmoins, la position de l’intron 3 est décalée de 3pb par rapport à celle de l’intron 3 des Vertébrés.

> Chez les deux espèces de nématode *Caenorhabditis*, bien que six introns existent, seules les positions des introns 1 et 5 des Vertébrés sont conservées.

> Chez l'abeille *Apis mellifera*, comme pour les *Caenorhabditis*, seules les positions des introns 1 et 5 des Vertébrés sont conservées. Sa séquence étant incomplète en 3', nous n'avons pu l'inclure pour la reconstruction de notre arbre phylogénétique. Chez les autres Insectes, *B. mori*, *A. gambiae* et les drosophiles, seule la position de l'intron 1 des Vertébrés est conservée.

En conclusion, dans les génomes de toutes ces espèces, la position de l'intron 1 est conservée sauf pour les Urochordés du genre *Ciona* qui ont des génomes très compacts (environ 116,7MB pour *Ciona intestinalis* contre, par exemple, 400MB chez le poisson *Takifugu rubripes*, 137MB chez *Drosophila melanogaster* ou encore 2,9GB chez *Homo sapiens*). Leur gène *Pofut1* a sans doute secondairement perdu ses séquences introniques. Sur la base de cette comparaison de l'organisation exon/intron, nous pouvons proposer que le gène *Pofut1* présent chez l'ancêtre des Bilatériens était constitué d'au moins deux exons et un intron en position 1.

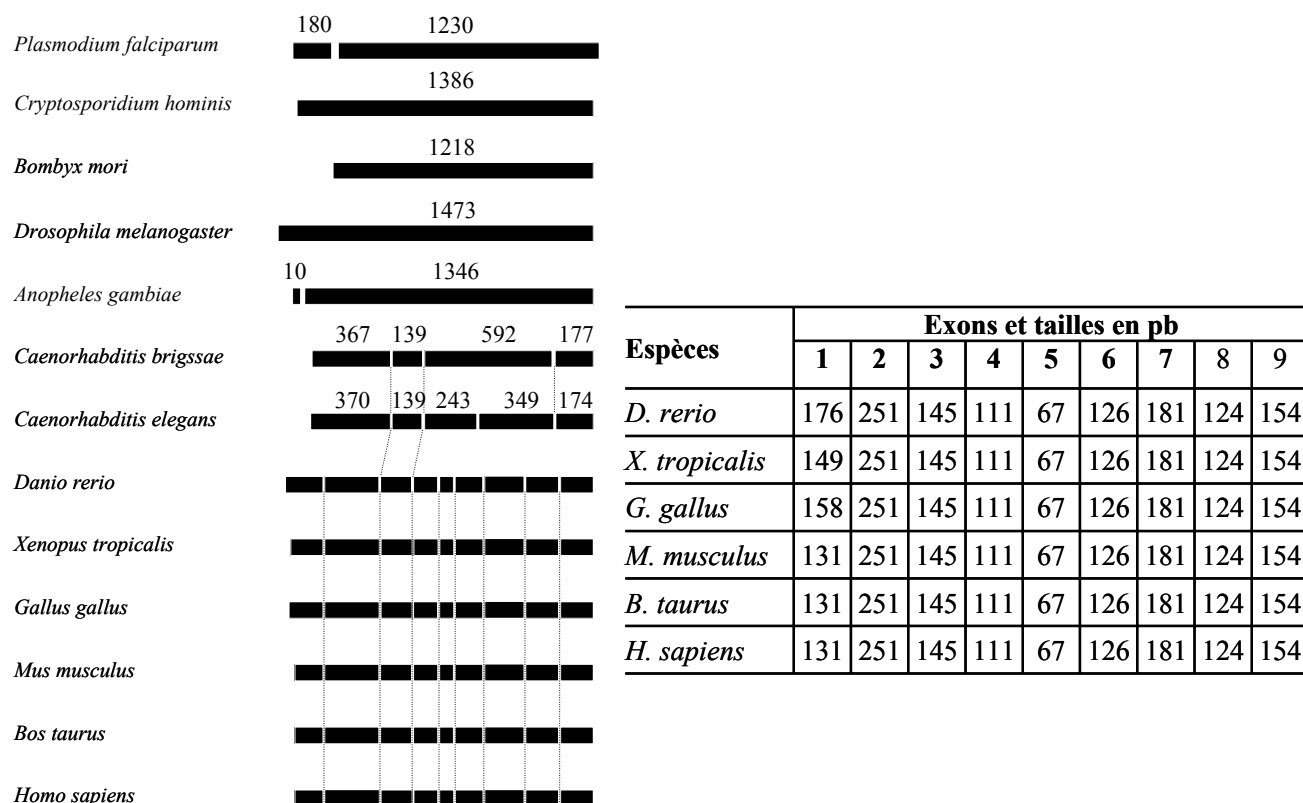


Figure 50 : Organisations exon/intron des CDS des gènes *Pofut2* chez différentes espèces animales.

Pour les Vertébrés, les tailles des exons homologues sont indiquées dans le tableau. Celles des invertébrés sont indiquées au dessus de chaque structure.

- Pour *Pofut2* les organisations génomiques révèlent huit introns à des positions strictement conservés chez les Vertébrés.

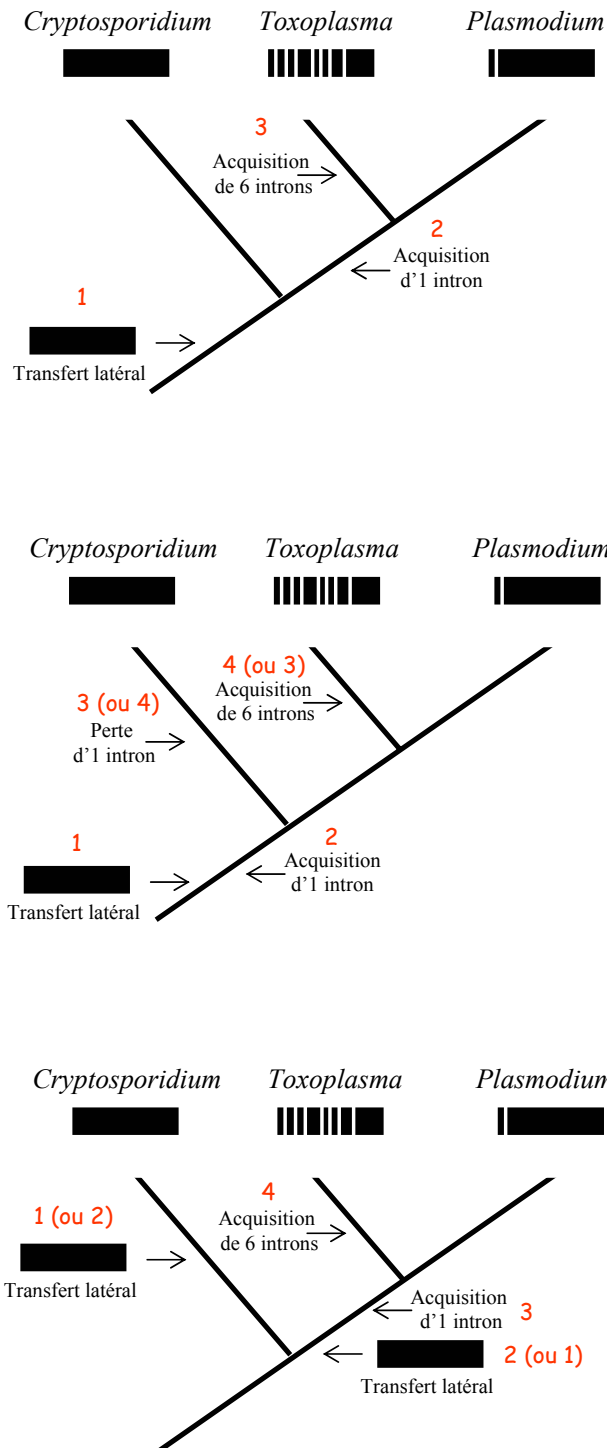
> Chez les nématodes *Caenorhabditis*, les positions des introns 2 et 3 sont conservées; cependant, *C. elegans* possède un intron supplémentaire, sans doute acquis après la séparation d'avec *C. briggsae*.

> Chez les Insectes *B. mori* et *D. melanogaster*, le gène *Pofut2* est monoexonique, alors que chez *A. mellifera*, il comporte un intron mais à une position non conservée par rapport à celle du premier intron des Vertébrés. Cet intron résulterait d'une acquisition secondaire propre à l'espèce.

Ainsi, le gène *Pofut2* chez l'ancêtre des Bilatériens devait sans doute posséder les introns 2 et 3 à ces mêmes positions. Ceux-ci auraient ensuite été perdus chez les Insectes, probablement par un mécanisme de transcription inverse et d'intégration dans le génome suite à une recombinaison homologe au même locus (Hu et Leung, 2006). De nouvelles séquences de *Pofut2* chez des espèces variées, notamment de protostomiens, devraient permettre de tester cette hypothèse.

En effectuant cette étude, nous avons observé que deux genres de protozoaires appartenant au phylum des Apicomplexés, *Cryptosporidium* et *Plasmodium*, possèdent chacun une enzyme homologue à *Pofut2*, respectivement de 461aa et 394aa. Récemment, nous avons identifié *in silico* la présence de *Pofut2* chez *Toxoplasma gondii*. Les membres de l'embranchement des Apicomplexés qu'on appelait autrefois Sporozoaires, sont tous des parasites unicellulaires d'Animaux; certains causent de graves maladies chez l'Homme, la cryptosporidiose lors d'une infection par *Cryptosporidium parvum*, la toxoplasmose dans le cas de *Toxoplasma gondii* et le paludisme pour *Plasmodium falciparum*. Chez *Cryptosporidium* et *Plasmodium*, *Pofut2* est codée respectivement par un et deux exons (Fig.50.). Chez *Toxoplasma gondii*, la protéine fait 478aa (www.toxoDB.org, séquence (TgGLEAN_4271)) et est codée par huit exons. Les séquences présentes chez ces organismes sont largement divergentes entre elles et par rapport à celles des Vertébrés. *Pofut2* de *C. parvum* et *P. falciparum* partagent 19,2% d'identité; celle de *T. gondii* partage 19,6% et 13,3% d'identité respectivement avec les séquences de *C. parvum* et *P. falciparum*. A titre d'exemple, l'enzyme *Pofut2* bovine partage 15,1% et 23% d'identité respectivement avec celle de *P. falciparum* et *C. parvum*. Nous n'avons pu mettre en évidence la présence de *Pofut2* chez d'autres espèces d'Apicomplexés (comme par exemple chez le genre *Eimeria*

responsable de la coccidiose, entre autres, chez les oiseaux et les bovins). Nous pouvons envisager que le gène *Pofut2* ait été acquis par ces parasites intracellulaires grâce à un transfert latéral lors de l'infection d'un hôte. La figure 51 résume les différentes hypothèses permettant d'expliquer la présence et l'évolution du gène *Pofut2* chez ces organismes.



Hypothèse 1

L'ancêtre commun des trois genres d'Apicomplexés a acquis par transfert latéral le gène *Pofut2* grâce à un mécanisme de rétrotranscription et insertion dans le génome. L'ancêtre commun des *Plasmodium* et des *Toxoplasma* a ensuite acquis un intron en position 1. Après la spéciation de ces derniers, l'ancêtre des *Toxoplasma* a acquis six introns.

Hypothèse 2

L'ancêtre commun des trois genres d'Apicomplexés a acquis par transfert latéral le gène *Pofut2* par rétrotranscription puis a acquis un intron en position 1. L'ancêtre des *Cryptosporidium* l'a ensuite perdu; leur génome est en effet beaucoup plus compact. L'ancêtre des *Toxoplasma* a ensuite acquis six introns

Hypothèse 3

L'ancêtre des *Cryptosporidium* et l'ancêtre commun des *Toxoplasma* et des *Plasmodium* ont acquis de façon indépendante le gène *Pofut2* à partir de transferts latéraux consécutifs à des infections d'hôtes différents. L'ancêtre commun des *Toxoplasma* et des *Plasmodium* a acquis un intron en position 1. L'ancêtre des *Toxoplasma* a ensuite acquis six introns.

Figure 51 : Représentation schématique des différentes hypothèses permettant d'expliquer la présence et l'évolution du gène *Pofut2* chez trois genres de protozoaires du phylum des Apicomplexés.

Les chiffres rouges indiquent le nombre d'évènements nécessaires à l'acquisition et l'évolution de ces gènes. L'hypothèse la plus parcimonieuse est l'hypothèse 1 qui ne fait intervenir que 3 évènements évolutifs contre 4 pour les deux autres hypothèses.

Le génome de ces Apicomplexés ne semble pas contenir de gène orthologue au gène *Pofut1*. Pourtant, chez *C. parvum* de nombreuses protéines extracellulaires possèdent des

domaines EGF et TSR (Templeton *et al.*, 2004), dont certains sont potentiellement *O*-fucosylables (Tableau 4). Cette constatation nous a amené à penser que la seule enzyme de *O*-fucosylation présente chez cette espèce, à savoir Pofut2, pouvait potentiellement avoir la fonction de fucosyler les deux types de domaines. Chez cette espèce, le gène *Pofut2* étant monoexonique, nous avons pu amplifier le CDS par PCR en utilisant de l'ADN génomique comme matrice (gracieusement fourni par Guan Zhu et Jan Keithly, Division of Infectious Diseases, Wadsworth Center, New York State Department of Health, Albany, New York, USA) et des amorces spécifiques choisies de part et d'autre de la séquence codante. L'amplification ainsi obtenue a ensuite été clonée dans le vecteur d'expression eucaryote pcDNA/TOPO3.1. L'expression hétérologue en cellules COS-1 de l'enzyme correspondante a permis de réaliser *in vitro* des tests fonctionnels qui n'ont révélé aucune activité *O*-fucosyltransférase sur les domaines TSR et EGF, utilisés pour la caractérisation des enzymes Pofut bovines. Cependant, des analyses par western blot n'ont pu mettre en évidence la présence de la protéine hétérologue, d'une masse estimée à 50 KDa (non glycosylée et sans le peptide signal), dans les extraits protéiques des cellules COS-1 transfectées. Malgré tout, nous pensons que l'enzyme Pofut2 de *C. parvum* a bien été exprimée, mais elle est tout simplement trop divergente dans sa séquence en acides aminés pour être reconnue par notre anticorps anti-Pofut2 bovin et pour fucosyler des motifs EGF ou TSR humains.

Tableau 4 : Séquences de domaines EGF et TSR *O*-fucosylables présents dans des protéines (non caractérisées) de *Cryptosporidium parvum*.

Les cystéines conservées participant à la formation des ponts disulfures sont en gras; le site potentiel de *O*-fucosylation est en gras, rouge et souligné.

Protéines	N° d'accès	Type	Séquences
Cgd3_1860	EAK89110	EGF n°1 sur 9	C ¹ LMPNYC ² SKEA T C ³ KSDIVGQQT ⁴ YIDYLNALPR C ⁴ Q C ⁵ LPGYIGDGR T KGTGC ⁶ Q
Cgd2_1590	EAK88814	EGF n°5 sur 5	C ¹ EENNP C ² DLSIS S C ³ INTPGSYIC ⁴ DC ⁵ FKSAGYMTSPIN N KIC ⁶ T
Cgd1_3500	EAK88310	TSR n°3 sur 6	VGSWSEWSD C ¹ ST S C ² GEGNRIR T REITK P PLNGD D SK C ³ PELIE K ES C ⁴ N K DC ⁵ PHV Q C ⁶
Cgd6_780	EAK90179	TSR n°1 sur 3	YSQWSSWLP C ¹ TS T C ² NGG T TYRWRYPLNEVMAGGLE C ³ TITSQ I S A C ⁴ NSD V EC ⁵ L P C ⁶

Une recherche dans les banques d'EST nous a également permis de trouver des séquences partielles codant potentiellement Pofut1 (CN775567 et CN774562) et Pofut2 (CN776950 et DN245882) chez un Hydrozoaire du phylum des Cnidaires, *Hydra magnipapillata*. Chez cette même espèce, nous avons également trouvé deux EST (DN815234 et CX831243) codant potentiellement des protéines ou des portions de protéines potentiellement *O*-fucosylables, notamment les domaines EGF 11 et 12 d'une protéine orthologue au récepteur Notch1. Ceci peut être corrélé à la mise en évidence chez l'hydre, de molécules impliquées dans la voie de signalisation par Notch, comme par exemple CBF1 et Numb (Voir chapitre 3 de l'exposé bibliographique) (Steele, 2002). De même, nous trouvons deux EST (DN246105 et DN246354) codant potentiellement une protéine homologue à la protéine CC6 (complement component 6) murine et possédant des domaines TSR *O*-fucosylables (Tableau 5). Les gènes *Pofut1* et *Pofut2* ainsi que les cibles des deux enzymes sont donc largement répandus chez les Animaux. La démonstration de leur existence reste à réaliser pour le phylum des Eponges.

Tableau 5 : Séquences de domaines EGF et TSR *O*-fucosylables chez *Hydra magnipapillata*.

Les cystéines conservées participant à la formation des ponts disulfures sont en gras; le site potentiel de *O*-fucosylation est en gras, rouge et souligné.

N° EST	Type	Séquences
DN815234	EGF n°11 de Notch	DIDEC ¹ STQPC ² ANNGS ³ IDIVNGFLC ⁴ RC ⁵ IDGF RGSNC ⁶ SIN
DN815234	EGF n°12 de Notch	SINIDEC ¹ DPSPC ² LHNS ³ VDQINGFQC ⁴ EC ⁵ SP GYFGLRC ⁶ ETEI
DN246105	TSR 1 de CC6	YSEWSSYTEC ¹ SATC ² GFGLMERERTC ³ TNPAPVKG GKDC ⁴ SSLGPMSEKPC ⁵ NLREC ⁶ P
DN246354	TSR 2 de CC6	ISEWTFDFGPC ¹ SKSC ² GGGISVTRTRSC ³ TNPSPLNG EKNC ⁴ SESIQESKDC ⁵ STDPC ⁶ P

1.2.5 Synthèse et conclusion

Comme dans les génomes de toutes les espèces étudiées jusqu'à présent, les gènes *Pofut1* et *Pofut2* ne sont présents qu'en une seule copie dans le génome de *Bos taurus* (Figure 3, Publication 1). L'analyse phylogénétique ainsi que des comparaisons de séquences montrent que les gènes *Pofut1* et *Pofut2* sont largement présents dans les génomes Animaux. Les organisations de chacun des deux gènes sont conservées chez les Vertébrés. En effet, si on ne tient compte que des exons codant la forme active de chaque enzyme, les positions introniques sont conservées pour *Pofut1* et *Pofut2* chez tous les Vertébrés (Figure 7A et B publication 1, Fig. 50 et 51.). L'analyse phylogénétique nous permet de proposer que les gènes de *O*-fucosyltransférases présents chez l'ancêtre des Bilatériens étaient constitués d'au moins deux exons et un intron en position 1 pour *Pofut1* et d'au moins trois exons et deux introns aux positions 2 et 3 pour *Pofut2*.

L'organisation génomique du gène *Pofut1* bovin est différente de celle précédemment déterminée chez l'Homme et la souris. En effet, le gène bovin que nous avons localisé *in silico* sur le chromosome 13, dans une région homologue au chromosome 20 humain et 2 murin, est plus long car il possède deux exons supplémentaires. Ceux-ci sont localisés environ 16 Kb après l'exon homologue des derniers exons humains et murins et semblent être une spécificité bovine. En effet, aucune séquence homologue n'a pu être mise en évidence dans les génomes d'autres espèces. Par rapport au gène murin, les gènes *Pofut1* humain et bovin présentent un exon supplémentaire en position 5. Néanmoins, les séquences de cet exon ne sont pas les mêmes chez les deux espèces. En effet, la séquence humaine homologue de l'exon 5 bovin est retrouvée dans l'intron 6.

Mise à part la taille des introns, plus petites dans le gène bovin, l'organisation génomique de *Pofut2* est la même que chez l'Homme et la souris. Pour l'instant, la localisation chromosomique du gène *pofut2* bovin n'est pas connue.

Dans l'espèce bovine, nous avons mis en évidence pour la première fois, cinq transcrits pour chacun des deux gènes et analysé leurs expressions tissulaires. Les tissus provenant d'embryons possèdent au moins trois variants *Pofut1* et les cinq variants *Pofut2*. Chez l'adulte, chaque tissu possède au moins un variant *Pofut1* et au moins quatre des cinq variants *Pofut2*. Un nombre plus important de transcrits différents pour chacun des gènes est donc observé chez l'embryon. Pour *Pofut1*, ce résultat est à corrélérer avec une activité intense du récepteur Notch lors du développement embryonnaire (Shi et Stanley, 2006). Il nous reste

cependant à déterminer si ces variants transcriptionnels sont traduits *in vivo*. Les protéines codées par les variants *Pofut1b* à *Pofut1e* et *Pofut2c* à *Pofut2e* n'ont pas d'activité *O*-fucosyltransférase *in vitro*, du moins sur les accepteurs que nous utilisons. Néanmoins, elles peuvent avoir des rôles importants dans la cellule. Il a été montré récemment que l'enzyme Pofut1 possédait une activité chaperonne vis-à-vis de Notch et ceci, indépendamment de son activité *O*-fucosyltransférase (Okajima *et al.*, 2005). Il est envisageable que les protéines issues de ces variants chez le bovin soient impliquées dans de tels mécanismes ou aient au sein des cellules des rôles de lectines reconnaissant des *O*-fucoses et modulant ainsi les interactions entre protéines *O*-fucosylées. Si toutefois ces transcrits ne sont pas traduits *in vivo*, ils peuvent avoir un rôle dans la régulation post-transcriptionnelle et donc contrôler le niveau d'expression des protéines actives (Lareau *et al.*, 2004); ce point est discuté dans la partie 2 des résultats. Le variant *Pofut2b* coderait une protéine dotée d'une faible activité *O*-fucosyltransférase 2 *in vitro*. Cependant, nous ne savons pas si cette protéine est traduite *in vivo*, et si c'est le cas, si elle possède, dans ce contexte, une activité *O*-fucosyltransférase significative.

Les anticorps anti-Pofut1 et anti-Pofut2 développés et testés dans le cadre de cette étude vont nous permettre de répondre à certaines de ces questions, en déterminant les profils d'expression des *O*-fucosyltransférases dans les tissus bovins. Ce point est abordé pour Pofut1 dans la seconde publication, et dans les résultats complémentaires, pour Pofut2. En effet, le second article est consacré à l'étude de la protéine Pofut1 bovine. Son expression tissulaire et ses caractéristiques fonctionnelles, notamment sa glycosylation, y sont détaillées.

Alignement des séquences protéiques de Pofut1 des différentes espèces choisies pour l'étude phylogénétique.

Les séquences sont alignées par rapport à la séquence de *Bos taurus* qui sert ici de référence. Les insertions et délétions sont symbolisées par * et les acides aminés identiques à la séquence de référence par -. Sous l'alignement, sont repérés par (*) les positions où l'acide aminé est strictement conservé. Les portions de l'alignement surlignées en gris ne sont pas prises en compte pour l'analyse phylogénétique (alignement délicat du aux degrés de divergence entre séquences, insertions spécifiques, séquences partielles). Les motifs I, II et III communs avec les α 2- et α 6-fucosyltransférases sont respectivement surlignés en bleu, rose et vert. Les sites potentiels de N-glycosylation sont surlignés en jaune. Les acides aminés définissant le peptide signal sont notés en gras et rouge, les motifs DXD sont notés en jaune et surlignés en bleu et les motifs KDEL-like sont notés en blanc, surlignés en rouge. Les positions des introns chez le bovin sont repérées par une flèche (\downarrow). Les numéros d'accès des séquences dans Genbank sont référencés dans la publication 1.

<i>Bos taurus</i>	MG*****	*****	AAAW*****	APSLPPRVS	LLLLLLPLP*	***GLPVG**	30
<i>Sus scrofa</i>	-----	*****	*****	-S-P---L--	-----*	***--A***	29
<i>Rattus norvegicus</i>	-----	*****	*****	---H-LL-A-	F-----L--L	PLR-RAG-**	34
<i>Mus musculus</i>	-----	*****	*****	---PH-LL-A-	F-----L--L	**R-RSA-**	32
<i>Canis familiaris</i>	-----	*****	*****	-S-P--L---	-----S----	*****--**	30
<i>Homo sapiens</i>	-----	*****	*****	-R***-LS--	F-----*	***-M-A-**	27
<i>Pan troglodytes</i>	-----	*****	*****	-R***-LS--	F-----*	***-M-A-**	27
<i>Pongo pygmaeus</i>	-----	*****	*****	---**LLS--	F-----*	***-M-A-**	27
<i>Xenopus tropicalis</i>	-RG*****	*****	*****	*****-W	-----SWV*	***RCERP**	20
<i>Xenopus laevis</i>	-ERG*****	*****	*****	*****-W	-F-----FV*	***-CEGL**	20
<i>Gallus gallus</i>	-VLL*****	***SGAALL	L**A*****	TV**TV**	*****	***AA-	19
<i>Danio rerio</i>	-AARPKH**	LKLFIVTYLH	T**S*****	LI**FL**	*****	***N-FID	28
<i>Oryzias latipes</i>	-ALKQGSITT	LSVLSLLPSL	LFO-C*****	LS**VT**	*****	***EGL**	32
<i>Takifugu rubripes</i>	-V*****	ICLYLLVFLH	L**C*****	QS**FS**	*****	***NGL**	21
<i>Tetraodon nigroviridis</i>	-V*****	ICLYLLVPH	L**C*****	QS**SK**	*****	***KEL**	21
<i>Ciona intestinalis</i>	-R*****	LRFYILTFVL	H**I*****	LT**FS**	*****	***GAA**	21
<i>Ciona savignyi</i>	-H*****	LSICILPLLL	R**T*****	LV**FA**	*****	***DGV**	21
<i>Strongylocentrotus purpuratus</i>	-RLL*****	LSLSSLVLSL	TLSLCEDHLD	SD**FQ**	*****	***S***	29
<i>Caenorhabditis briggsae</i>	-KILR*V**	LVVYSIYSIY	F**A*****	DA**VK**	*****	***NGA**	25
<i>Caenorhabditis elegans</i>	-RVSK*V**	LTLASFISVC	S**Y*****	SE**AK**	*****	***S***	23
<i>Aedes aegypti</i>	*****	*****	*****	*****	*****	*****	0
<i>Anopheles gambiae</i>	-*****	**YRLVGLLL	T**F*****	LT**FA**	*****	***HSI**	18
<i>Bombyx mori</i>	-ITV*****	***ILMVLL	T**T*****	TS**V***	*****	***KSS**	19
<i>Drosophila melanogaster</i>	-QWLKMK**	LRFVNLILL	ISSTC*****	-Q**LGG*	*****	*****	27
<i>Drosophila pseudoobscura</i>	-QWLKMK**	FPFIYLICTV	T----*	-N**LE--*	*****	*****	27
<i>Drosophila yakuba</i>	-QWLKMK**	LRFVNLILL	ISS-C*****	-Q**LGG*	*****	*****	27
<i>Onchocerca volvulus</i>	*****	**FLLFYLF	T**L*****	LWIGNFTS*	*****	*****	17

						Intron1	
						↓	
<i>Bos taurus</i>	****S****	*****	*****	*****WDP	AGYLLYCPCM	GRFGNQADHF	54
<i>Sus scrofa</i>	***V-****	*****	*****	*****---	-----	-----	54
<i>Rattus norvegicus</i>	***-*****	*****	*****	*****--L	-----	-----	58
<i>Mus musculus</i>	***-*****	*****	*****	*****--L	-----	-----	56
<i>Canis familiaris</i>	***-*****	*****	*****	*****---	-----	-----	54
<i>Homo sapiens</i>	***-*****	*****	*****	*****---	-----	-----	51
<i>Pan troglodytes</i>	***-*****	*****	*****	*****---	-----	-----	51
<i>Pongo pygmaeus</i>	***-*****	*****	*****	*****---	-----	-----	51
<i>Xenopus tropicalis</i>	*****	*****	*****	*****--S	S--I-----	-----	43
<i>Xenopus laevis</i>	*****	*****	*****	*****YNS	N--IV-----	-----	43
<i>Gallus gallus</i>	***-*****	*****	*****	*****--T	-----E--	-----	43
<i>Danio rerio</i>	ATDVT****	*****	*****	*****--E	N--I-----	-----V--	56
<i>Oryzias latipes</i>	***H****	*****	*****	*****--E	N--V-----	-----	56
<i>Takifugu rubripes</i>	***R****	*****	*****	*****--Q	N--V-----	-----	45
<i>Tetraodon nigroviridis</i>	***-*****	*****	*****	*****--Q	N--V-----	-----	45
<i>Ciona intestinalis</i>	***E****	*****	*****	*****I-E	N--IA-----	-----	45
<i>Ciona savignyi</i>	***E****	*****	*****	*****Y-K	N--VI-----	-----V--	45
<i>Strongylocentrotus purpuratus</i>	***NEGLV	VTDEREDGVF	SAGGAGLEVS	QAAPKLE---	N--FTF-----	-----E-Y	84
<i>Caenorhabditis briggsae</i>	***KNGT**	*****	*****	*****--	N--IVF-----	-----I-Q-	51
<i>Caenorhabditis elegans</i>	***NET**	*****	*****	*****--	N--IVF-----	-----V-Q-	48
<i>Aedes aegypti</i>	*****	*****	*****	*****	*****	-----	10
<i>Anopheles gambiae</i>	***T****	*****	*****	*****V-E	H--IM-----	-----	42
<i>Bombyx mori</i>	***N****	*****	*****	*****PEK	K--IA-----	-----	43
<i>Drosophila melanogaster</i>	*****	*****	*****	*****--	N--T-----	-----	49
<i>Drosophila pseudoobscura</i>	*****	*****	*****	*****--	N--T-----	-----	49
<i>Drosophila yakuba</i>	*****	*****	*****	*****--	N--T-----	-----	49
<i>Onchocerca volvulus</i>	*****	*****	*****	*****E	T--VV-----	-----ME--	38

** * * * * *

						Intron 2	
						↓	
<i>Bos taurus</i>	LGSLAFAKLL	NRTLAVPPWI	EYQHKKPPFT	NV***HVSQ	KYFKLE*PLQ	AYHRVISLED	110
<i>Sus scrofa</i>	-----	-----	-----	-L***	-----*	V-----	110
<i>Rattus norvegicus</i>	-----	-----	-----	-L***	-----*	-----	114
<i>Mus musculus</i>	-----	-----	-----	-L***	-----*	-----V--	112
<i>Canis familiaris</i>	-----	-----	-----R----	-L***	-----*	-----	110
<i>Homo sapiens</i>	-----	-----	-----	-L***	-----*	-----	107
<i>Pan troglodytes</i>	-----	-----	-----	-L***	-----*	-----	107
<i>Pongo pygmaeus</i>	-----	-----	-----	-L***	-----*	-----	107
<i>Xenopus tropicalis</i>	-----MV	---V----	V-N--R--Y-	---***--P-E	EF-Q--*-R	QF-H---M-Q	99
<i>Xenopus laevis</i>	-----MV	---V----	V-N--R--Y-	---***--P-E	EF-Q--*-R	QF-Q---M-E	99
<i>Gallus gallus</i>	--A---RA-	-----	--R--R--Y-	-L***--P-E	E-----*	Q---L---Q	99
<i>Danio rerio</i>	-----M-	-----	V-R--S----	---***--P-S	E--Q--*-K	Q---V----	112
<i>Oryzias latipes</i>	-----M-	-----	V-R--I--Y-	---***--P-S	EF-Q-D*V-S	-----V----	112
<i>Takifugu rubripes</i>	-----M-	-----	V-R--V--Y-	---***-IP-R	DF-Q--*A-S	-----V----	101
<i>Tetraodon nigroviridis</i>	-----M-	-----	V-R--A--YA	---***--P-G	EF-Q--*A-S	-----V----	101
<i>Ciona intestinalis</i>	---S---K-	---II----	VHGYRGYGD*	DS***FIP-S	RW-QVN*A-K	P----EMKE	100
<i>Ciona savignyi</i>	---S---RK-	---V----	THKYGRYDG*	DS***FPP-N	HW--VD*T-K	S---I-EM--	100
<i>Strongylocentrotus purpuratus</i>	---V---G-A-	---IIL--FR	T-*****	*I***NIPFT	DW-QF-PM-E	F-P-L-TM-T	132
<i>Caenorhabditis briggsae</i>	---V-S---S-	D---VL-NF-	-FN**H-*E*	TK***MIPFE	VL-QVG*TVA	K-T--VTMHE	103
<i>Caenorhabditis elegans</i>	---V---A-	D---VL-NF-	-FK**H-*E*	TK***MIPFE	FL-QVG*TVA	K-T--VTMQE	100
<i>Aedes aegypti</i>	---A-G--HG-	---VL---V	--RKGAAN**	S-***Q-PFD	T--QVA*--T	E-T---TM-N	64
<i>Anopheles gambiae</i>	---FG---A-	---VL---V	--RKGEV*R*	SI***Q-PFD	T--QVG*---	-F----TM-N	96
<i>Bombyx mori</i>	---V---S-AV	---IL---	--RYGE-*R*	SI***Q-PFD	T--NVD*K-L	EFAN--TM--	97
<i>Drosophila melanogaster</i>	-----A-	---IL---V	--RRGEL*R*	SR***Q-PFN	T--EV---K	E-----TMA-	103
<i>Drosophila pseudoobscura</i>	-----G-	---IL---V	--RKGEM*R*	SR***Q-PFN	T--EV---K	D-----LMS-	103
<i>Drosophila yakuba</i>	-----A-	---IL---V	--RRGEL*R*	SR***Q-PFN	T--EV---K	E-----TMA-	103
<i>Onchocerca volvulus</i>	---M---S-	---VL--F-	---**H-*E*	KKQAIM-DFD	T--LV---R	NF-KM-VMR-	92

** * * * *

Intron 3

				↓				
<i>Bos taurus</i>	FMEKLAPTHW	PPEKRVAYCF	EVAA*****	QRSPDKKTC	MKEGNPFGPF	WDQFHVSE	NK	164
<i>Sus scrofa</i>	-----	---R-----	-----*****	-----	-----	-----		164
<i>Rattus norvegicus</i>	-----F--	-----	-----*****	-----	-----	-----		168
<i>Mus musculus</i>	---N---S--	-----	-----*****	-----	-----	-----		166
<i>Canis familiaris</i>	-----	---R-----	-----*****	-----	-----	-----		164
<i>Homo sapiens</i>	-----	-----	-----*****	-----	-----	-----		161
<i>Pan troglodytes</i>	-----	-----	-----*****	-----	-----	-----		161
<i>Pongo pygmaeus</i>	-----	-----	-----*****	-----	-----	-----		161
<i>Xenopus tropicalis</i>	---Q---SVL-	--SR-----	TA--*****	-----S--	--D-----	--H---N-IQ		153
<i>Xenopus laevis</i>	---Q---VL-	--T--L----	TA--*****	-----S--	--D-----	--H---D-TQ		153
<i>Gallus gallus</i>	---Q---Q--	--GR-----	A--*****	---A--S--	--D-----	---K-D-D-		153
<i>Danio rerio</i>	---H---K--	-HGQ-I----	S--*****	H--Q--S--	--D-----	--HIG-D-DR		166
<i>Oryzias latipes</i>	-----	---GQ-R----	T--*****	---A--S--	--D-----	--YVG-N-D-		166
<i>Takifugu rubripes</i>	---T---RY-	-AGQ-R----	T--*****	--TA--S--	--D-----	--YVN-D-DE		155
<i>Tetraodon nigroviridis</i>	---A-----	-SGQ-R----	A--*****	---E--S--	-----	---HVN-D-DQ		155
<i>Ciona intestinalis</i>	--LEI--SI-	L---YI--H-	-STF*****	K--E--S--	A-S-----	--N-KID-PA		154
<i>Ciona savignyi</i>	--TN---SI-	--N--KI--H-	-I-F*****	S--D--S--	A-S-----A-	--N-K-E-IA		154
<i>Strongylocentrotus purpuratus</i>	---E---KY-	-ENE-RG--W	LYLNSDAQ**	*****-E	--R-----	--G-GIN-AE		182
<i>Caenorhabditis briggsae</i>	-TK-IMSK-F	ACT*****	*PRKAIYD**	**TDAEPG-H	A-----Y	---IG---VG		151
<i>Caenorhabditis elegans</i>	-TK-IM---F	VGT*****	*PRQAIYD**	**KSAEPG-H	S-----Y	---ID---VG		148
<i>Aedes aegypti</i>	--SN---SL-	---E-ISF-Y	TERMG*LD**	**GSTGGG-N	A-S-----	--T-KID-V-		119
<i>Anopheles gambiae</i>	--KT---SL-	--AE-ISF-Y	TERMG*LD**	**GSTGHG-N	A-S-----	--T-E-E-AG		151
<i>Bombyx mori</i>	--REI---V-	-I---SF-Y	TQRTGEI***	*****QNS-N	A-S-----	--TYNID--R		149
<i>Drosophila melanogaster</i>	--WH--DDI-	-ESE--SF-Y	KERYS*LQQE	KND--PN-H	A-D-----	--T--ID-VR		162
<i>Drosophila pseudoobscura</i>	--WH--DDI-	-ESE--SF-Y	MERKS*LQQE	KNN--QPN-H	A-D-----	--TYNID-V-		162
<i>Drosophila yakuba</i>	--WH--DDI-	-ES--SF-Y	KERYS*LQQE	KND--PN-H	A-D-----	--T--ID-VR		162
<i>Onchocerca volvulus</i>	-IKEI--NI-	-LD--K-F-W	QPRQSILD**	**TRKSSG-H	A-----SI	LGL-KCN-YW		148
	*			*	* * * * *		*	

Intron 4

				↓				
<i>Bos taurus</i>	SELFAG*IS*	*FSASY****	*K****DQW*	*****	IQRF	SPEEHPVLAL	PGAPAQFFVL	205
<i>Sus scrofa</i>	-----*--*	*-----****	*-----*--*	*****	---	--K-----	-----	205
<i>Rattus norvegicus</i>	---T-*--*	*-----****	*-----E--*	*****	---	P-K-----	-----	209
<i>Mus musculus</i>	---T-*--*	*-----****	*-----E--*	*****	T---	PAK-----	-----	207
<i>Canis familiaris</i>	K--S-*--*	*-----****	*-----*--*	*****	---	--K-----	-----	205
<i>Homo sapiens</i>	---T-*--*	*-----****	*R****E--*	*****	S---	--K-----	-----	202
<i>Pan troglodytes</i>	---T-*--*	*-----****	*R****E--*	*****	S---	--K-----	-----	202
<i>Pongo pygmaeus</i>	---T-*--*	*-----****	*R****E--*	*****	S---	--K-----	-----	202
<i>Xenopus tropicalis</i>	---D-*--T*	*---Y-****	*-----V-*	*****	NS--	P-SQ---I--	-----	194
<i>Xenopus laevis</i>	---D-*--T*	*---Y-****	*-----V-*	*****	S--	P-SQ---I--	-----	194
<i>Gallus gallus</i>	---K-*--*	*---SA-****	*R****H--*	*****	---	P-S-----	-----	194
<i>Danio rerio</i>	-V--G-*L*	*---SY-****	*Q****PH-*	*****	-K--	P-K-----	-----	207
<i>Oryzias latipes</i>	-V--G-*L*	*---SYH****	*Q****L--*	*****	MKK-	P-S-----	-----M	207
<i>Takifugu rubripes</i>	-V--G-*Y*	*---Y-****	*Q****P--*	*****	MKK-	P-SQ---V-	-----S	196
<i>Tetraodon nigroviridis</i>	-V--G-*Y*	*---S-****	*Q****P--*	*****	MKK-	P-SQ---I--	-----L--Q	196
<i>Ciona intestinalis</i>	--G-P*HLH*	*-HS****S*	*-----KE--*	*****	DSVY	PATI-----F	M----SY--N	195
<i>Ciona savignyi</i>	--G-P-NLN*	*YHS****S*	*-----TS--*	*****	DHAY	PS-T-L---F	M----AY--K	195
<i>Strongylocentrotus purpuratus</i>	EIK-EIFSHA	DL-NTQDMEK	I-****EH--*	*****	DRS-	PVKD-----	K----P--MR	231
<i>Caenorhabditis briggsae</i>	D-Y-GE*-PG	G-ET-QMGS*	*R****KK--*	*****	LEK-	G---F---F	-S---A--SK	197
<i>Caenorhabditis elegans</i>	D-Y-GD*-PG	G-DLNQM***	***GSRKK--*	*****	LEK-	PS--Y---F	SS---P--SK	194
<i>Aedes aegypti</i>	--F-GPKLN*	*YDVYH****	*HD*MANK--*	*****	AE-Y	PSDKW---I-F	T---S---IQ	164
<i>Anopheles gambiae</i>	--F-GPKLN*	*YDIHH****	*GDSMARR--*	*****	SE-Y	PAAKW---F	T---A---Q	197
<i>Bombyx mori</i>	--FY-P*LN*	*YD-QN****	*EK*ILE--*	*****	KI-Y	PVDKW---F	T---S---Q	193
<i>Drosophila melanogaster</i>	--FY-P*LN*	*-DVHHS***	***NEAAK--*	*****	QTKY	PA-SY---F	T---S---Q	206
<i>Drosophila pseudoobscura</i>	--FYGP*LN*	*-DVHHS***	***SMATK--*	*****	QSKY	PADSW---F	T---S---Q	206
<i>Drosophila yakuba</i>	--FY-P*LN*	*-DVHHS***	***NEAAK--*	*****	QAKY	PA-SY---F	T---S---Q	206
<i>Onchocerca volvulus</i>	RW-LRF*RI*	*-LKHI****	*L****LA-Q	ACQKNGMRN-	QLI-IQF--V	HFP--A-SR-		196
			*				**	

Intron 5
↓

<i>Bos taurus</i>	EEHRPLQKYM	VWSD	***MV	RTGEAQIH	LI*RPYVGIH	LRIGSDWKNA	CAMLKDG	TAG	261
<i>Sus scrofa</i>	-----	-----	***--	-----R-	-*-	-----	-----	-----	261
<i>Rattus norvegicus</i>	---E---	---	***--	-----ST-	-V*	-----	-----	-----	265
<i>Mus musculus</i>	---E---	---	***--	-----L-SS-	-V*	-----	-----	-----	263
<i>Canis familiaris</i>	-----V	---	***--	-----H-LT-	-*	-----	-----	-----	261
<i>Homo sapiens</i>	-----	---	***--	K-----	-V*	-----	-----	-----	258
<i>Pan troglodytes</i>	-----	---	***--	K-----	-V*	-----	-----	-----	258
<i>Pongo pygmaeus</i>	-----	---	***--	KV-----	-V*	-----	-----	-----	258
<i>Xenopus tropicalis</i>	---S--R-V	---EK***	I-	-A--E--QSL	-V*	---V-	-----	-K-----	250
<i>Xenopus laevis</i>	---S--R-I	---EK***	I-	-E--E--RSL	-V*	-----	-----D-	-K-----	250
<i>Gallus gallus</i>	-----H--	---	***--	KK---Y--SL	-V*	-----	-----	-N-----	250
<i>Danio rerio</i>	---IG--QFV	---K***	I-	QE--GH-RNL	-N*	-----	-----Q--	-KL--T-DT-	263
<i>Oryzias latipes</i>	---VG--RLV	---R***	--	TE--EH-A-M	-T*	---L---	-----QS-	-RL-ES-D-	263
<i>Takifugu rubripes</i>	---VG--R--	---EK***	--	EE-DGH-EKL	-V*	-----	-----QT-	-G--DS-DM-	252
<i>Tetraodon nigroviridis</i>	G--VG--R--	---EK***	--	QE-DQH-RTL	-*	---V-	---L-V--QT-	-G--AS-EV-	252
<i>Ciona intestinalis</i>	-KDKQ--R-V	Q-TK-***	LE	NSA-KL-DLK	-K*K	-----	---N-M--K-	-DHV-N-L**	249
<i>Ciona savignyi</i>	AGD-S----V	E-K-***	IK	NAANEV-KNK	-Q*	-----	---N-M--R-	-EHV---I**	249
<i>Strongylocentrotus purpuratus</i>	ASN-VI-N-F	K-NEG***	LF	AEAKHHMDHL	FNGE	-----	---T-I--ERS	-ETVEGMR**	286
<i>Caenorhabditis briggsae</i>	GKVWSI---L	R--SR***	IT	EQAKKY-TNN	-Q*K-F-AV-	-----	---NDA--VRV	-DHIDLEK**	251
<i>Caenorhabditis elegans</i>	GKVWSI---L	R--SR***	IT	EQAKKF-S-N	-A*K-F-AV-	-----	---NDA--VRV	-EHIDTT-**	248
<i>Aedes aegypti</i>	I-N-E-H--L	K--VT***	IE	NAARDF-RNS	-PKGAFM---	-----	---N-I--IR-	-DHV-ESS**	219
<i>Anopheles gambiae</i>	Q-NLL-HR-L	R--VK***	YE	TGAKQF-RDT	-PRGAFI-V-	-----	---N-I--VR-	-EHV-ESS**	252
<i>Bombyx mori</i>	K-NVDM--L	K--ND***	IE	HRSKLF-KK	MSGGGFL---	-----	---N-Q--VK-	-QHV--SP**	248
<i>Drosophila melanogaster</i>	L-NCK--R-L	Q--QR***	YR	EASKDF-REQ	-PRGAFL---	-----	---N-I--VR-	-EHV--SQ**	261
<i>Drosophila pseudoobscura</i>	L-N-E---H-	Q--AK***	YR	DAARDF-RKV	-PRGAF---	-----	---N-I--VR-	-EHV--SQ**	261
<i>Drosophila yakuba</i>	L-NCK--R-L	H--QR***	YR	DASKDF-REQ	-PRGAFL---	-----	---N-I--VR-	-EHV--SQ**	261
<i>Onchocerca volvulus</i>	WPDHF-FCRN	LSIGPKRAI-		KEANKYVAES	-K*--FI---	-----	---NDV--E-L	-R-D-KTKLR	255

* * * * *

<i>Bos taurus</i>	AH**FMAS	PQ	CVG***YSRH	TTAPLMTMC	LPDLKEIRRA	LKLWVTAL**	*****NAQSV	309	
<i>Sus scrofa</i>	---	---	***--	-A	-----	---	*****S---	309	
<i>Rattus norvegicus</i>	S--*	---	***--S	-AT--T---	-----Q--	V---R--**	*****D-R--	313	
<i>Mus musculus</i>	S--*	---	***--S	-AT--T---	-----Q--	VT---R--**	*****--R--	311	
<i>Canis familiaris</i>	S--*	---	***--S	-A	-----	V---E--**	*****---	309	
<i>Homo sapiens</i>	S--*	---	***--S	-A	-----	V---RS--**	*****D---	306	
<i>Pan troglodytes</i>	S--*	---	***--S	-A	-----	V---RS--**	*****D---	306	
<i>Pongo pygmaeus</i>	S--*	---	***-G-S	-A	-----	V---RS--**	*****D---	306	
<i>Xenopus tropicalis</i>	S--*	---	***-DKY	RA	-----E--	-T---ERS**	*****K-R--	298	
<i>Xenopus laevis</i>	S--*V--A-	---C***-D-Y	RAG	-----D--	-----M--	-T---ERS**	*****K-R--	298	
<i>Gallus gallus</i>	S--*	---	***-D-S	-AV--D--	-----K--	--V--KKT**	*****V-K-I	298	
<i>Danio rerio</i>	P--*	---	***-D-Q	-AL--N--	-----T--	V---KNT**	*****G---	311	
<i>Oryzias latipes</i>	P--*	---	***--Q	-AL--A--	-----A--	V---KKT**	*****S-R--	311	
<i>Takifugu rubripes</i>	P--*	---	***-E-Q	-AL--A--	-----G--L--	V-V--KKT**	*****S-R--	300	
<i>Tetraodon nigroviridis</i>	P--*	---	***-E-Q	-AL--PA--	-----Q--L--	V---EKT**	*****SSR--	300	
<i>Ciona intestinalis</i>	*DYP--S---	-T-***-Q-S	-SKRF-YE--	Y-S-NTVKSH	VLSSL-KES**	*****-SN-L	298		
<i>Ciona savignyi</i>	*DYP--S---	-T-***-K**	**GRF-YE--	F-SDNTVKTH	VLHVAK-I**	*****G-N-L	294		
<i>Strongylocentrotus purpuratus</i>	*Q**M-----	-HKVG**Q**	**E-V-KA--	Y-TKEMVLEE	TRAA-EKY**	*****KTKH-	330		
<i>Caenorhabditis briggsae</i>	*NRPLF--E-	-L-EGHHQ**	**GQ---E--	I-NKQQ-VDM	IVEK-GSI**	*****G-K--	299		
<i>Caenorhabditis elegans</i>	*NRPLF--E-	-L-EGHHL**	**GT--KEI-	S-SKQQ-LEQ	IVEK-GSI**	*****G-K--	296		
<i>Aedes aegypti</i>	*N**LFS---	-L-YRNEK**	**GS--LE--	--SKEI-I-Q	I-RQIKNHKE	AYKNNEIK--	272		
<i>Anopheles gambiae</i>	*N**LFS---	-L-YRNEH**	**GQ--AEL-	M-GRDT-I-Q	--RQIKLHRE	AAPDNPIRA-	305		
<i>Bombyx mori</i>	*M**LF-A--	---YRNER**	**G---S--	--KPS--LKR	-FSRQIKRVI	K*KLSDIKIYI	300		
<i>Drosophila melanogaster</i>	*--*LF----	-L-YKNER**	**GA-YPEL-	M-SKEA-I-Q	--RTIKNVRQ	TQPDNEIK--	314		
<i>Drosophila pseudoobscura</i>	*--*LF----	-L-YKNER**	**GS-YPEL-	MQTKE-S-V-Q	--RTIKNVRQ	TQPNNEIK--	314		
<i>Drosophila yakuba</i>	*--*LF----	-L-YKNER**	**GA-YPEL-	M-SKEA-I-Q	--RTIKNVRQ	TQPDNEIK--	314		
<i>Onchocerca volvulus</i>	NYLVQR	FVLV	***	*****A	IN--VE--	--S-ET-IKD	IMKKASD--**	*****KVR--	298

*

Intron 6

↓

<i>Bos taurus</i>	YIATDSESYL	PEIQQLFKG*	**KVKVVS*	KPEVAQIDLY	ILGQADHFIG	NCVSSFTAFV	365
<i>Sus scrofa</i>	-----V	S-----E*	**-----*	-----	-----	-----	365
<i>Rattus norvegicus</i>	-----V	S-----D*	**--R-----*	-----	-----	-----	367
<i>Mus musculus</i>	-----V	-----E*	**-----*	-----D	-----	-----	365
<i>Canis familiaris</i>	-V-----V	--L-----*	**-----*	-----V	-----	-----	362
<i>Homo sapiens</i>	-V-----V	--L-----*	**-----*	-----V	-----	-----	362
<i>Pan troglodytes</i>	-V-----V	--L-----*	**-----N*	-----V	-----	-----	362
<i>Pongo pygmaeus</i>	-----T-HT	E--AFLGD*	**--R--C*	Q----M---	-----	-----	354
<i>Xenopus tropicalis</i>	-----P--T	A--AFLGN*	**E--R--C*	Q-----	---A---	-----	354
<i>Xenopus laevis</i>	-----P--T	K--F-Q*	**--I-----*	Q-----	-----S	-----	354
<i>Gallus gallus</i>	-----HT	T--K-----*	**-----*	Q-DT--V---	-----	-----S	367
<i>Danio rerio</i>	-----VAHR	ED-EK--E*	**-----*	H-DW--T---	-----	-----S	367
<i>Oryzias latipes</i>	-----HS	GD-E--N*	**-----*	R--L--M---	---K---	-----S	356
<i>Takifugu rubripes</i>	-----K-HS	GD-EK--R*	**-----*	Q--S--T---	V-RK---	---W--S	356
<i>Tetraodon nigroviridis</i>	F----AS--E	N-FNNM--EA	GLD-SI-K*	ENDELM--A	L-IES-R---	S-G--V---	357
<i>Ciona intestinalis</i>	F----AK-FE	E-FKTW--DA	GLNIDIK*	QKDELN--A	V-IE--T---	S-G--V---	353
<i>Ciona savignyi</i>	F----KF--Q	K-LKEM-GD*	**E--HH*	D-HMP---	I-----	SEF---	S
<i>Strongylocentrotus purpuratus</i>	FV-S-KDHMI	D--NEAL-PY	E**IEAHRQ*	EVDDMYTS-A	-M-RS-L---	----T-SHI-	356
<i>Caenorhabditis briggsae</i>	FV-S-KDHMI	D--NEAL-PY	E**IEAHRQ*	E-DDMYTS-A	-M-R--L-V-	----T-SHI-	353
<i>Caenorhabditis elegans</i>	FV-S--NHMI	HDLNEAL-RM	D**IT--K*	PENNPHL--A	---MSN---	----YS---	329
<i>Aedes aegypti</i>	FV-S--NHMI	G-LNDAL-RM	D**T--RH*	PDGNPHL--A	---R-N---	--I--YS---	362
<i>Anopheles gambiae</i>	FV-S--NHMI	E-LNAGLRRP	E**I-ILKFQ	PSSNPHL--A	---NY---	----YS---	358
<i>Bombyx mori</i>	FV-S--NHMI	G-LNTALSRM	G**IS-HK*	PEDDPYL--A	---SN---	--I--YS--E	371
<i>Drosophila melanogaster</i>	FV-S--ANHMI	G-LNAALSRM	G**IS-HR*	PEDDPYL--A	---SN---	--I--YS---	371
<i>Drosophila pseudoobscura</i>	FV-S--NHMI	G-LNTALSRM	G**IS-HK*	PEDDPYL--A	---SN---	--I--YS---	371
<i>Drosophila yakuba</i>	FVSS-RNH-I	D-LK-KLASL	Q**IV-ERR*	Y--NLH-N-A	--SI-----	---TLS--I	355

* * * * *

<i>Bos taurus</i>	KFRERDLHGRP	***SSFFGMD	RPPQIRDEF*	*****	***	391
<i>Sus scrofa</i>	-----Q--L	***-----	-----S	*****	***	391
<i>Rattus norvegicus</i>	-----Q	***-----	-----S	*****	***	393
<i>Mus musculus</i>	-----Q--L	***-----	K-----	*****	***	391
<i>Canis familiaris</i>	-----Q--L	***-----	-----K	*****	***	388
<i>Homo sapiens</i>	-----Q--L	***-----	-----K	*****	***	388
<i>Pan troglodytes</i>	-----Q--L	***-----	-----K	*****	***	388
<i>Pongo pygmaeus</i>	-----V--L	***-----	S-GAVQ--L*	*****	***	380
<i>Xenopus tropicalis</i>	-----D--V	***-----	S-GVVQ--L*	*****	***	380
<i>Xenopus laevis</i>	-----V--K	***-----	H--R-----	*****	***	380
<i>Gallus gallus</i>	-----NI---	***-----L-	Y-GKKNRNEE	I*****	***	395
<i>Danio rerio</i>	-----HT--S	***-----	--RKANAKEE	I*****	***	395
<i>Oryzias latipes</i>	-----Q--VQ-L	***-----	T-GN--VREE	I*****	***	384
<i>Takifugu rubripes</i>	-----V--L	***-----	T-GRVNAREE	I*****	***	384
<i>Tetraodon nigroviridis</i>	S-K-EVL---	***NEY--AS	DKHNEA-RSE	I*****	***	385
<i>Ciona intestinalis</i>	S-K-KVLD--	***NE--NL	GNQNNA-HTE	I*****	***	381
<i>Ciona savignyi</i>	-----GN-K	***--T-WSIG	V*****	*****	***	404
<i>Strongylocentrotus purpuratus</i>	-----HT-QS	PRP-A---VR	SARRHI-L**	*****	***	384
<i>Caenorhabditis briggsae</i>	-----HA-QS	PRP-A---IR	AVKRHI-L**	*****	***	381
<i>Caenorhabditis elegans</i>	-----AK-F-	***--T-WAFP	A*EKHMSKSA	KAAAGGSVTH	EEL	368
<i>Aedes aegypti</i>	-----AN-F-	***--WAFP	P*EKPGAKVK	GG**GSSPQH	EEL	399
<i>Anopheles gambiae</i>	-----MK-L-	***--E-WSFP	Q*RKKNKHEE	I*****	***	385
<i>Bombyx mori</i>	-----Y--F-	***--Y-W-FP	K*EKD-KHTN	V*****	H EEL	402
<i>Drosophila melanogaster</i>	-----S--F-	***--Y-WSFP	K*EKD-QQSK	A*****	H EEL	402
<i>Drosophila pseudoobscura</i>	-----V--F-	***--Y-W-FP	K*EKD-KHSN	V*****	H EEL	402
<i>Drosophila yakuba</i>	Y-Q-KYTSNV	ARE----AQN	YFIENK--L*	*****	***	384

* *

Alignement des séquences protéiques de Pofut2 des différentes espèces choisies pour l'étude phylogénétique.

Les séquences sont alignées par rapport à la séquence de *Bos taurus* qui sert ici de référence. Les insertions et délétions sont symbolisées par * et les acides aminés identiques à la séquence de référence, par -. Sous l'alignement, sont repérés par (*) les positions où l'acide aminé est strictement conservé. Les portions de l'alignement surlignées en gris ne sont pas prises en compte pour l'analyse phylogénétique (alignement délicat du aux degrés de divergence entre séquences, insertions spécifiques, séquences partielles). Les motifs I, II et III communs avec les α 2- et α 6-fucosyltransférases sont respectivement surlignés en bleu, rose et vert. Les sites potentiels de N-glycosylation sont surlignés en jaune. Les acides aminés définissant le peptide signal sont notés en gras et rouge, les motifs DXD sont notés en jaune et surlignés en bleu. Les positions des introns chez le bovin sont repérées par une flèche (\downarrow). Les numéros d'accès des séquences dans Genbank sont référencés dans la publication 1.

<i>Bos taurus</i>	MAALA***** VVFLLLWAAS WPSAS***** *****AF GEE**FW***	27
<i>Rattus norvegicus</i>	----S***** --C---A--- -RPV-***** *****-S ----*--***	27
<i>Mus musculus</i>	----S***** --C---A--- -RPV-***** *****-S ----*--***	27
<i>Homo sapiens</i>	--T-S***** F-----G-V- --P--***** *****-S -Q-*--***	27
<i>Pan troglodytes</i>	--T-S***** F-----G-V- --P--***** *****-S -Q-*--***	27
<i>Gallus gallus</i>	---HHPG*** ***** **--MLLLL VLATL***** *****-- LPPAAA***Q	28
<i>Xenopus tropicalis</i>	-AV***** ***** *****-- -RPRLLLLG VCLSLLPR-C -D-GQ-S***	32
<i>Danio rerio</i>	--TNRGTAVS VHLHHILGYI T--IFVILS- KS-SA***** *****SE DDA**R***	42
<i>Oryzias latipes</i>	***** ***** *****FFLSI FAAFA***** *****-V KS-NV-S***	19
<i>Takifugu rubripes</i>	***** ***** ***** ***** ***** *****	0
<i>Ciona intestinalis</i>	-SNC-L*** *****YF ILN-VFILL- YNY-V***** *****CE DAK*****	28
<i>Ciona savignyi</i>	--KII***** ***** FC-I-ITLL- YNLVL***** *****CD NNNSK*****	27
<i>Caenorhabditis briggsae</i>	-H***** ***** **--PFLLLV LG-SV***** *****-L RSKVDQHHKV	26
<i>Caenorhabditis elegans</i>	-H***** *****FF PIQ--VLFFA EKI-F***** *****-E NSDQ*T***V	27
<i>Anopheles gambiae</i>	-NR***** ***** *****TVY- LLLLA***** *****FS QIVHT*****	19
<i>Bombyx mori</i>	-K***** ***** *LISTIFFNL ILINA***** *****-I NT-NGYC**	25
<i>Drosophila melanogaster</i>	-RGSWPRLG* ***** *****FPALLL LLHLL***** *****TG SDAAVR***N	29
<i>Drosophila pseudoobscura</i>	-HKMRWRGS G***** *LR---VLLL L--SI***** *****-I ADRPAT***R	34
<i>Drosophila yakuba</i>	-RGSWPRLG* ***** *****-PALLL LLHLL***** *****TV ADANFR***N	29
<i>Plasmodium berghei</i>	-I***** ***** YNLISISLY- LIIIL***** *****TD IYAKT*****	24
<i>Plasmodium falciparum</i>	-K***** ***** F IIV---FFFF KVIDR***** *****VI CVTPQK*****	26
<i>Plasmodium vivax</i>	-KGR-HIWA L***** ***** *****LL***** *****-C LPPFRF***N	22
<i>Cryptosporidium hominis</i>	-INFLYSILF SCILN***** **VR--FFIF IVVNA***** *****-N NSDQ*C**I	36
<i>Cryptosporidium parvum</i>	-TNFLYPILL SCILN***** **IR--FFIF TVANA***** *****-N NSDQ*C**I	36

Intron 1



<i>Bos taurus</i>	*****PGQ	SAADIL*SGA	*****AS	RRRYLLYDVN	PPEGFN**LR	59
<i>Rattus norvegicus</i>	*****---	-----*---	*****---	-----	-----**--	59
<i>Mus musculus</i>	*****---	-----*---	*****---	-----	-----**--	59
<i>Homo sapiens</i>	*****---	-----*---	*****---	-----	-----**--	59
<i>Pan troglodytes</i>	*****---	-----*---	*****---	-----	-----**--	59
<i>Gallus gallus</i>	SLPTALRA-H	A-TLSLAAA-	*****P	HT-----	-----**--	68
<i>Xenopus tropicalis</i>	*****--H	--SSLQVGT-	*****F	E-----	-----**--	65
<i>Danio rerio</i>	*****A--	--MFSPV*AA-	*****SR	DV-----	-----**--	74
<i>Oryzias latipes</i>	*****A-H	KSTAPI*A*-	*****K	DL-----	-----**--	50
<i>Takifugu rubripes</i>	*****---	*****---	*****---	***-----	-----**--	15
<i>Ciona intestinalis</i>	*****---	*****---	*****P-G	NK-----C-	-G-----**--	49
<i>Ciona savignyi</i>	*****V	*****---	*****P-	NK-----C-	-G-----**--	48
<i>Caenorhabditis briggsae</i>	SHVDSNKYSI	*****---	*****A	EKKF-----	FG-----**--	56
<i>Caenorhabditis elegans</i>	SRVDSNRYSV	*****---	*****A	EKKF-----	FG-----**--	57
<i>Anopheles gambiae</i>	*****TSAHL	RELCEKRDIF	FEPYRCQLVD	LDLLDDP***	*VL-----S-----**--	68
<i>Bombyx mori</i>	*****VSK	DSCAETKNT*	*****S-	DKKW-F----	-----**--	57
<i>Drosophila melanogaster</i>	GTAKREIGDS	RGSSGTCVKG	FLQEILPLPA	TCPPEVLGMR	GAV-I-----IS-----**--	87
<i>Drosophila pseudoobscura</i>	NRTESHRTAV	RSNSNVCKSR	QFWQTIPLVD	TCSPGILPLR	GTV-I-----IS-----**--	92
<i>Drosophila yakuba</i>	GTARPANGAS	RGSSATCAKR	VLQEILPLPE	TCPPEVLGLR	GTV-I-----IS-----**--	87
<i>Plasmodium berghei</i>	*****SSL**	**ICNKNDIY	LNDKFY****	*****FLK	KKK-I-----IG-----YS-Q	64
<i>Plasmodium falciparum</i>	*****L**	**ICLKEDVY	LGDDFF****	*****FLK	-KK-IM----IG-----**--Q	62
<i>Plasmodium vivax</i>	LDKDVSS-*	**VCRTDDVY	TGDAFY****	*****PFK	KKK-V----H IG-----**--Q	65
<i>Cryptosporidium hominis</i>	NILNSWDLR-	GMENKCPNWV	V*****	*****DVCWII--K	DG-----**--Q	75
<i>Cryptosporidium parvum</i>	NILNRWDLR-	GMENKCPNWV	V*****	*****DV-WII--K	DG-----**--Q	75

** **** *

<i>Bos taurus</i>	RDVYIRIASL	LKTLKLT*E*	***EW**VLV	LPPWGRL*YH	WQSPDIHQVR	IPWSDFFDLP	111
<i>Rattus norvegicus</i>	-----V---	-----*--	***---	-----*	-----	---E-----	111
<i>Mus musculus</i>	-----V---	-----*--	***---	-----*	-----	---E-----	111
<i>Homo sapiens</i>	-----V---	-----*--	***---	-----*	-----	---E-----	111
<i>Pan troglodytes</i>	-----V---	-----*--	***---	-----*	-----	---E-----	111
<i>Gallus gallus</i>	-----M---	-----S*--	***N--*	-----*	-----L---	---E-----	120
<i>Xenopus tropicalis</i>	-----V---	-----S*S*	***N--*	-----*	-----	---AE----D	117
<i>Danio rerio</i>	-----M-T-	-----R-E*G*	***D--*	-----*	-----E---L-	---GE--SIT	126
<i>Oryzias latipes</i>	-----M---	-----V---R-EGG*	***D--*	-----*	-----S--D-T-	---GE--SIT	103
<i>Takifugu rubripes</i>	-----M---	-----V---T-DGD*	***D--*	-----*	-----Y-L-	---GE--S-T	68
<i>Ciona intestinalis</i>	---M-V-N-	V-S-RESKSC	KGDD--*	-----IS-	-KEHGME-SL	---M--NVK	107
<i>Ciona savignyi</i>	---M--N-	V-N-RERK-C	KGDD--*	-----IG-	-KER-LE-SM	---T---K	106
<i>Caenorhabditis briggsae</i>	---M-V-NT	VRL-RESG-*	***N--*I-	-----*H-	---**RM--A	LS-RL---E	106
<i>Caenorhabditis elegans</i>	---M-V-NT	VRS-RDSG-*	***NY**I-	-----*H-	-K***RME-A	LS-RL---E	107
<i>Anopheles gambiae</i>	-----L-VF	VQF-RTQRGY	RR**T**R--	-----SS-*V-	-R-GN-D-QQ	LL-GN----	123
<i>Bombyx mori</i>	---M-F-I-	-SNEQ-RGKL	KE**F**G--	-----YK-*--	-KT*TKKREP	---GN---IE	111
<i>Drosophila melanogaster</i>	-----M-VF	VRR-QRRRRF	RH**V**R--	-----P--*--	-H-QGLQ-SG	L---H---A	142
<i>Drosophila pseudoobscura</i>	-----L-VF	VRR-RRRKR	QN**V**R--	-----FP--*--	-H-HHLK-ND	VA-GH---D	147
<i>Drosophila yakuba</i>	-----V-VF	VRR-QRRRRF	RQ**V**R--	-----P--*--	-H-QGLQ-AG	L---H---D	142
<i>Plasmodium berghei</i>	KE-LY-MSLV	IYQ-N-RDKE	YI**Y**Y--	-----CY-*S-	*--*NKRHDN	LQ-NI-LNMN	117
<i>Plasmodium falciparum</i>	KEIFY-LSLV	IYN-N-KDKI	NI**Y**Y--	-----CYV*T-	-NI**RKGN	LR-EF--NTD	115
<i>Plasmodium vivax</i>	KE-LY-V-LA	VYY-NQEERT	HV**H**Y--	-----CYV*T-	-GR*ERTNA-	-K--I--N-K	119
<i>Cryptosporidium hominis</i>	-N-FD-M-LV	VSS-NEIMYK	KNEKFA--I	---FCNI*A-	-TY*NSKR*-	L---T--SIK	132
<i>Cryptosporidium parvum</i>	-N-FD-M-LV	VSS-NEIMYK	KNEKFA--I	---FCNI*A-	-TY*NSKR*-	L---T--SIK	132

* * * * *

Intron 2



<i>Bos taurus</i>	SLNRNIPVIE	YEQFIAESGG	*****PF	***IDQVYIL	QGY**TEGWK	EG****	AWE*	154
<i>Rattus norvegicus</i>	--K-----	-----	*****--	***-----V-	---**A----	---****T---		154
<i>Mus musculus</i>	--K-----	-----	*****--	***-----V-	---**A----	---****T---		154
<i>Homo sapiens</i>	--K-----	-----	*****--	***-----V-	-S**A----	---****T---		154
<i>Pan troglodytes</i>	--K-----	-----	*****--	***-----V-	-S**A----	---****T---		154
<i>Gallus gallus</i>	--K-----	---L-----	*****--	***-E-I-V-	---**E----	---****T---		163
<i>Xenopus tropicalis</i>	--YQ-----	--D-L-----	*****--	***-E--V-	---**A----	---****T---		160
<i>Danio rerio</i>	--QA-----	--E-----	*****--	***---ILV-	-S**A---T	D****K---		169
<i>Oryzias latipes</i>	--QA-V--V-	--E---N--	*****--	***---LV-	-N**A---T	D****K---		146
<i>Takifugu rubripes</i>	--QV-V----	--E---N--	*****--	***---LV-	-N**A---T	D****K---		111
<i>Ciona intestinalis</i>	-MSGFV----	--DY-E-V-E	*****AV	***---ELWY-	-N**A---S	S****T---		150
<i>Ciona savignyi</i>	---EFV--M-	--DY-E-I-E	*****AS	***---ELWY-	-N**A---T	S****V---		149
<i>Caenorhabditis briggsae</i>	----F---M-	F-D-LD-NRP	*****	***---IY-	-H**E---G	T****EYVR		146
<i>Caenorhabditis elegans</i>	----F-----	F-D-LD-NRP	*****	***---IY-	-H**A---G	T****EYVR		147
<i>Anopheles gambiae</i>	-MRLYTD-LD	MDE-F--Y-R	LHGP GATTVV	***L-E--K-	KHF**EQMFE	N****VFVD		174
<i>Bombyx mori</i>	--KSFA--V-	LYEIFEK-LQ	KQ*****LI	***--TI-V-	-N**DNPFE	N****LF-D		156
<i>Drosophila melanogaster</i>	--R-YA--LD	--E-L--QRL	F-NPGAPLVH	***VGHAFR-	-H**EVMLE	Q****IFRD		193
<i>Drosophila pseudoobscura</i>	--R-YA--LD	-SE-L--YRM	F-MPAPPYVH	***-S--FR-	MH**EIMLE	Q****IFRD		198
<i>Drosophila yakuba</i>	--R-YA--LD	-DE-L--QRI	F-NPGAPLVH	***VAHAFR-	-H**KVMLE	H****IFRD		193
<i>Plasmodium berghei</i>	I IKNV--I--	FSDYK KLY-*	*D*****VT	DFI-SFK-L-	DSNAQKA***	*****		156
<i>Plasmodium falciparum</i>	IMKKV--I--	--EYEKLY-*	*N*****YS	DIM-NSK---	DN**K-***	*****		152
<i>Plasmodium vivax</i>	A-QNV--M-	-AEYEQF-*	*P*****HT	DYILSYRH-I	*****	**GEWPK***		156
<i>Cryptosporidium hominis</i>	**QNGL-TL-	-DIYKRLT-I	*****VN	***-GAEVVG	LNF**DWN-	SL****	QSST	173
<i>Cryptosporidium parvum</i>	**QNGL-TL-	-DIYKRLT-I	*****VN	***-GAEVVG	LNF**DWN-	SF****	QSST	173

<i>Bos taurus</i>	KIDSR****	**PCIE****	PPL**YS***	*****	*****	*****	*****	168
<i>Rattus norvegicus</i>	-V-E-*****	**--D****	-L-***--	*****	*****	*****	*****	168
<i>Mus musculus</i>	-V-A-*****	**--D****	-L-***--	*****	*****	*****	*****	168
<i>Homo sapiens</i>	-V-E-*****	**--D****	QL-***--	*****	*****	*****	*****	168
<i>Pan troglodytes</i>	-V-E-*****	**--D****	QL-***--	*****	*****	*****	*****	168
<i>Gallus gallus</i>	---E-*****	**--D****	QLM***--	*****	*****	*****	*****	177
<i>Xenopus tropicalis</i>	-V-Q-*****	**--D****	QLM***--	*****	*****	*****	*****	174
<i>Danio rerio</i>	-V-D-*****	**--D****	KLM***--	*****	*****	*****	*****	183
<i>Oryzias latipes</i>	---E-*****	**--FD****	KLM***--	*****	*****	*****	*****	160
<i>Takifugu rubripes</i>	-V-K-*****	**--D****	KLI***--	*****	*****	*****	*****	125
<i>Ciona intestinalis</i>	-MHE-*****	**D---****	N-V**E***	*****	*****	*****	*****	164
<i>Ciona savignyi</i>	-MHE-*****	**D--D****	S-V**E***	*****	*****	*****	*****	163
<i>Caenorhabditis briggsae</i>	-F-E-*****	**R-LP****	-ADSH-R***	*****	*****	*****	*****	162
<i>Caenorhabditis elegans</i>	-FEK-*****	**S-LP****	-AESH-K***	*****	*****	*****	*****	163
<i>Anopheles gambiae</i>	-FEEH*****	**V-PKGHDG	ALA*****	*****	*****	*****	*****	190
<i>Bombyx mori</i>	-WEVQ*****	**EECP****	*****	*****	*****	*****	*****	165
<i>Drosophila melanogaster</i>	-FERV*****	**TDKP****	CSEGS L-G**	*****	*****	*****	*****	210
<i>Drosophila pseudoobscura</i>	-YERV*****	**APKLDNNS	CNEYSTAG**	*****	*****	*****	*****	219
<i>Drosophila yakuba</i>	-FERV*****	**TDKP****	CSEGS L-G**	*****	*****	*****	*****	210
<i>Plasmodium berghei</i>	***KYHYHVL	PFDKCYIEDY	KFKIVC*KNC	DYKYSVTYSG	NCTNIKGNKT	LCYSDYIVTN		211
<i>Plasmodium falciparum</i>	***KSFLIL	PFEECNINVN	RFKQFC*KKC	EHKYNVLYSG	YCTTINTKQS	ECYSYMNISN		207
<i>Plasmodium vivax</i>	RG-KKSFQVL	KLDKCQVKG Y	KLKKNLRKNC	DHKYSVEYSG	KCTNVK GK KM	ECLEFFFFITS		216
<i>Cryptosporidium hominis</i>	SIEYY*****	**NSFK****	ISESRHLNSK	CF*****	*****	*****		187
<i>Cryptosporidium parvum</i>	SIEYY*****	**NSF-****	ISESRHLNSK	CF*****	*****	*****		187

Intron 3

↓

<i>Bos taurus</i>	*****	*****	*****QDKH	EYYRGWFWGY	EETRGLNVSC	LSVQGSASII	202
<i>Rattus norvegicus</i>	*****	*****	*****	---F---	---	-----V-	202
<i>Mus musculus</i>	*****	*****	*****	---	---	-----V	202
<i>Homo sapiens</i>	*****	*****	*****	---	---	-----V	202
<i>Pan troglodytes</i>	*****	*****	*****	---	---	-----V	202
<i>Gallus gallus</i>	*****	*****	*****K--	Q---	---	-----VV	211
<i>Xenopus tropicalis</i>	*****	*****	*****E--N	G---	---I--T	-----M	208
<i>Danio rerio</i>	*****	*****	*****K--Q	G---	---F---A--T	--A--H--L	217
<i>Oryzias latipes</i>	*****	*****	*****K--Q	G---	---AR--T	V-A--H--M	194
<i>Takifugu rubripes</i>	*****	*****	*****K--Q	GF---	---AR--T	I-A--H--M	159
<i>Ciona intestinalis</i>	*****	*****	*****T-SE	GKI---	---YAKTFK-	V---MSKLF	198
<i>Ciona savignyi</i>	*****	*****	*****T-TE	GKI---	---YAKKFK-	V---MSKMF	197
<i>Caenorhabditis briggsae</i>	*****	*****	*****KEND	SKW---YS-	-DVQAS-FK-	V-I--DSGTL	196
<i>Caenorhabditis elegans</i>	*****	*****	*****-VEE	FKWK---YS-	-DVYSR-FQ-	V-I--DSGTL	197
<i>Anopheles gambiae</i>	*****	*****	*****	*****YF--	GNFTADAF-	-HF---MLL	214
<i>Bombyx mori</i>	*****	*****	*****	YVPQ**Y--	GNITANEII-	VKF--KI-ML	193
<i>Drosophila melanogaster</i>	*****	*****	*****	***-PLLQQ	A-L-VGRFH-	VRF---GLL	236
<i>Drosophila pseudoobscura</i>	*****	*****	*****	***-PLLSQ	NLL-YGQYH-	VHF---GLL	245
<i>Drosophila yakuba</i>	*****	*****	*****	***-PLLQQ	A-L-VGRFH-	VRF---GLL	236
<i>Plasmodium berghei</i>	YLVNYVVHKL	FY*****	*****	*****	*****	*****	223
<i>Plasmodium falciparum</i>	YFITSILENL	F*****	*****	*****	*****	*****	218
<i>Plasmodium vivax</i>	HFVSTLLDI	F*****	*****	*****	*****	*****	227
<i>Cryptosporidium hominis</i>	*****	*****	*****	***FNNYVR	NGNVVYGGR-	EL-KIRGL**	218
<i>Cryptosporidium parvum</i>	*****	*****	*****	***FNNYVR	NGNVVYGGR-	EL-KIRGL**	218

Intron 4

↓

Intron 5

↓

<i>Bos taurus</i>	A*PVLRLN**	TSA*****	RS VM*****	***LDRAENL	LHDHYGGKEY	WDTRRSMVFA	243
<i>Rattus norvegicus</i>	-*---K--**	*****	---	*****	---	-----R-	243
<i>Mus musculus</i>	-*---K--**	*****	---	*****	---	-----R-	243
<i>Homo sapiens</i>	-*-L---**	*****	---	*****	---	-----	243
<i>Pan troglodytes</i>	-*-L---**	*****	---	*****	---	-----	243
<i>Gallus gallus</i>	-*-I--K--**	*****Q-	---	*****	---	-----D- -N-----	252
<i>Xenopus tropicalis</i>	-*---K--**	-T*****	---	*****	---	-----D-	249
<i>Danio rerio</i>	-*---QN--**	IT*****	T-	*****	***V---TV	---A--D-	258
<i>Oryzias latipes</i>	-*---Q--**	YTV*****	T-	*****	***--T-T-	---V-A--D-	235
<i>Takifugu rubripes</i>	-*---QK--**	ITV*****	T-	*****	***--T-	---A--D-	200
<i>Ciona intestinalis</i>	VK-LCGG--**	-T*****	K-	*****	***--G-M	---MF---R- -C-----	240
<i>Ciona savignyi</i>	VQ-ICGG--**	-T*****	K- L	*****	***--G-M	---MF---N- -C-----S	239
<i>Caenorhabditis briggsae</i>	K*DL-THSNF	S-M*****	T-	*****	***V---TI	--ENF-DVN- -KA----RYS	239
<i>Caenorhabditis elegans</i>	K*DL-KHSNF	SES*****	T- I	*****	***V---TI	--E--EVD- -KA----RYS	240
<i>Anopheles gambiae</i>	H*RL-E-Y**	RPGRPRAT-	Y -H*****	***VLN--IV	---LW-NVD-	-EA----R--	260
<i>Bombyx mori</i>	P*ELFQLH**	RND*****	KF I-	*****	***VAHG-IP	---N--DRK- -KC-K--K-N	234
<i>Drosophila melanogaster</i>	E*KL-REA**	IDE*****	DT AGPEDVDDMR	TYA-LS--TV	---W-DEHF	-QA----R--	288
<i>Drosophila pseudoobscura</i>	E*RL-REA**	ITE*****	DT AGAEDVDDMR	VYAILS--TV	---W-DEHF	-KA----R-S	297
<i>Drosophila yakuba</i>	E*KL-REA**	IAE*****	DT AGPEDADDMR	TYA-LS--TV	---W-DEHF	-QA----R--	288
<i>Plasmodium berghei</i>	*****I**	HNI*****	N- IL*****	***IKNSNVV	-VPFPNELFE	NNIEDILL-N	258
<i>Plasmodium falciparum</i>	*****L**	YNI*****	T- -L*****	***IKQST-I	-VPFVNELYQ	SNLEDILL-N	253
<i>Plasmodium vivax</i>	*****Q**	YD-*****	D- -L*****	***IKHGS-I	-VAFMNELVD	ANLEDVLPYS	262
<i>Cryptosporidium hominis</i>	***IC-SF**	FKL*****	MR FK*E*****	***VSEEIYK	ILESNRKEGK	LHNYLIKHTE	258
<i>Cryptosporidium parvum</i>	***IC-SF**	FKL*****	MR FK*E*****	***VSEEIYK	ILESNRKEGK	LHNYLIKHTE	258

Intron 6



<i>Bos taurus</i>	RHLRAVGDEF	RSRYLNSTDV	ADGIPFEEDW	TRMKVKLG**	*SSLGGPYLG	VHLRRKDFIW	300
<i>Rattus norvegicus</i>	K-----	--QH---A	--KMVL---	-----**	*-A-----	-----	300
<i>Mus musculus</i>	K-----	--QH---A	--KMAP---	-K-----**	*-A-----	-----	300
<i>Homo sapiens</i>	---E----	--H---D	--R---Q---	MK-----**	*-A-----	-----	300
<i>Pan troglodytes</i>	---E----	--H---D	-----Q---	-K-----**	*-A-----	-----	300
<i>Gallus gallus</i>	K---VA----	-NK--Q---E	--RTHYN---	-Q---T---**	*TA-----	-----	309
<i>Xenopus tropicalis</i>	K---V----	-ANF-H---E	--RTI-H---	-KV-A-PD**	*-A-----A	-----	306
<i>Danio rerio</i>	K---II--K-	-ANH---K	Q-QT-YN---	-Q--NP---**	*-AK-----	-----	314
<i>Oryzias latipes</i>	K---LI----	-AK---D	R-LTVYR---	---A-P---**	*-AK-----	-----	292
<i>Takifugu rubripes</i>	K---LI--N-	-ATH---I	S-STVYS---	---A---**	*-AK-----	-----	257
<i>Ciona intestinalis</i>	DA-IKTA-Q-	--K--D-D--	R-KTVM*PEN	WKDHKCTE**	*DAI----IA	A-V-----LY	296
<i>Ciona savignyi</i>	HK-IEIANN-	--KF-D-N-I	E-NTVM*PEN	WKNHKCN-**	*DAT----IA	A-V-----LY	295
<i>Caenorhabditis briggsae</i>	KD-VEAANV-	-DK--D-D-E	N-DTEM*VAD	W-KERAKR**	*TAK-----	I-W--R--LY	295
<i>Caenorhabditis elegans</i>	ND-VD-A-A-	-KK--D-D-K	R-KTKL*V-D	WTKEKPRR**	*TAI-----	I-W--R--LY	296
<i>Anopheles gambiae</i>	PP-VD-ATR-	-AE---S-R	--RTVPRAR-	-DERSRPR**	*KAR--E--C	A---A--LY	317
<i>Bombyx mori</i>	KK-TEIAFNY	IKEK--C***	*****	*****D**	*TE-CKN---	--W--Q--AK	271
<i>Drosophila melanogaster</i>	-R-EQ-AAD-	-RQA-DT--A	S*AGVQRPAM	WELERPKR**	*NAK--D--C	A---G--VR	344
<i>Drosophila pseudoobscura</i>	KY-SQ-AAD-	LRYAF*D-TD	TTAGVQRPAM	WE-ERPKR**	*DAR--DFIC	A---G--VR	353
<i>Drosophila yakuba</i>	-R-DQ-AAD-	-LHA-AT--A	S*VGVQRPAL	WELERPKR**	*DAK--D--C	A---G--VR	344
<i>Plasmodium berghei</i>	EK-IYN-NNY	I*****	*****	*****KE	*ILKNSN-IS	C---YN--KK	290
<i>Plasmodium falciparum</i>	NK-LSY-NNY	I*****	*****	*****S	*NILKTNH-IS	S---YT--KY	285
<i>Plasmodium vivax</i>	ED-INE--Q-	VEK*****	*****	*****	*NFKSSKN-IS	C---YT--RK	295
<i>Cryptosporidium hominis</i>	GI-VPWPW-L	SKYRVLDVIS	YNNNIRRYAN	LY-SANTF**	*FEEKK--IS	-----N--VF	315
<i>Cryptosporidium parvum</i>	GI-VPWPW-L	SKYRVLDVIS	YNNNIRRYAN	LY-SANTF**	*FDEKK-FIS	-----N--VF	315

* * *

Intron 7



<i>Bos taurus</i>	GHREDVPSLD	GAVRKIRSLM	KAHQ**LDKV	FVATDAVRKE	LEELRRLLE	MVR***FEP	354
<i>Rattus norvegicus</i>	-----E	---K-----	-T---**	-----I---	Q---K---	-----	354
<i>Mus musculus</i>	-----E	---K-----	-T---**	-----I---	Q---K---	-----	354
<i>Homo sapiens</i>	---Q----E	-----	-T-R**	-----	Y---KK---	-----	354
<i>Pan troglodytes</i>	---Q----E	-----	-T-R**	-----	Y---KK---	-----	354
<i>Gallus gallus</i>	-----H	--AK--H--L	-T-K**E--	-I-----ED-	I-L-KK-V--	-----	363
<i>Xenopus tropicalis</i>	-----R	R-ADE-H---	-KLK**K--	-----D-ED	--A--K-I--	-----A	360
<i>Danio rerio</i>	-----K	---K-----	-K-K**LQ-	-----DEE-	-AK-K-----	-----	369
<i>Oryzias latipes</i>	-----K	---KT-LN--	-K-M**E--	-I-----DGE-	-K--KAM--R	-----	346
<i>Takifugu rubripes</i>	-----K	---K-----	-K-K**---	-L-----DGG-	-K--K-----	-L-****-D-	311
<i>Ciona intestinalis</i>	ARK-F---IK	QVAKVLSEK-	-EYK**V--	-I-S-GTKE-	MK--KS---G	-FM***YK-	350
<i>Ciona savignyi</i>	ARQ-Y---IK	QVAKVLSEK-	-EYN**VN--	-I-S-GTKD-	M---KS---G	-IM***-K-	349
<i>Caenorhabditis briggsae</i>	ARHDQL--IK	-TAEILEK-C	EKLK**NRI	YL----PEE-	VN--KSYFGK	ELQ***VF*	348
<i>Caenorhabditis elegans</i>	AR-AQL-TIP	-TAKILQD-C	-KLD**Q-I	YL----PDQ-	VD--KA--NG	ELE***VY*	349
<i>Anopheles gambiae</i>	-RDKTT-TVQ	S-ALQ--AKL	LELG**KT-	---S-CT-M-	FYN-KNY-KR	FRV*VRFVPE	374
<i>Bombyx mori</i>	YRSK---TTS	-T-KQ-QNI-	QNKLPHVQT-	-I-----SA--	R-I-QNN-ET	LGYNVHYV-	331
<i>Drosophila melanogaster</i>	SRDATT-T-K	A-AQQVKQ-L	RGFN**MTT-	-L-----TPY-	-M--KE-FYR	FRL*VHFape	401
<i>Drosophila pseudoobscura</i>	SREATT-N-K	S-AQQ-KQ-L	R-FN**ATT-	-LS---TPYD-	-VQMKMNFYG	FRF*VHFQPE	410
<i>Drosophila yakuba</i>	AREATT-T-K	A-AQQ-KQ-L	R-FN**MTT-	-L-----TPY-	-M--KE-FYR	FRL*VHFape	401
<i>Plasmodium berghei</i>	ITSY---VQ	I-IL-LLYI-	FMNN**KE-I	-IS--EKKYV	QYIINKHFQK	FKH***-F*	343
<i>Plasmodium falciparum</i>	IS-YN--PIH	I-LL-LLYI-	FINN**CRII	-I-S-EKVEI	QKVINKDFHQ	YKK***HF*	338
<i>Plasmodium vivax</i>	ISTY--SPVG	ISLL-LLYI-	FLRK**STLI	--S--EKKEV	KKVIDSQF-Q	FKH***-F*	348
<i>Cryptosporidium hominis</i>	LRND-I-TF-	QVIDRLLQ-S	-DLK**TKR-	VIS--SNEY-	KN--LDIFLK	HNLDLHII**	372
<i>Cryptosporidium parvum</i>	LRND-I-TF-	QVIDRLLQ-S	NDLK**TKR-	VIS--SNEY-	KN--LDIFLK	HNLDLHII**	372

*

Intron 8



<i>Bos taurus</i>	TWEE*LELYK	DGGVAIVDQW	ICAHARFFIG	TSVSTPFSFRI	HDEREILGLD	*****	PKTTY	408
<i>Rattus norvegicus</i>	---*---	---I---	---	---	---	*****	---	408
<i>Mus musculus</i>	---*---	---I---	---	---	---	*****	---	408
<i>Homo sapiens</i>	---*---	---I---	---	---	---	*****	---	408
<i>Pan troglodytes</i>	---*---	---I---	---	---	---	*****	---	408
<i>Gallus gallus</i>	---*---	---MVI---	---	---	---	*****	---	417
<i>Xenopus tropicalis</i>	---*---	---I---	---	---	---	*****	---	414
<i>Danio rerio</i>	SL-D*---	---L---	---	---	---	*****	---	423
<i>Oryzias latipes</i>	SR-D*---	---L---	---	---	---	*****	---	400
<i>Takifugu rubripes</i>	SI-D*---	---L---	---	---	---	*****	---	365
<i>Ciona intestinalis</i>	-KT-*KSEF-	M---	---	---	---	*****	---	404
<i>Ciona savignyi</i>	SKA-*--NNF-	M---	---	---	---	*****	---	403
<i>Caenorhabditis briggsae</i>	RFTD*TQK-N	--QI-	---	---	---	*****	---	402
<i>Caenorhabditis elegans</i>	RFTD*TQKLN	--QI-	---	---	---	*****	---	403
<i>Anopheles gambiae</i>	SY-Q*RAEL-	---	---	---	---	*****	---	428
<i>Bombyx mori</i>	NIA-*VTEFD	--TI-	---	---	---	*****	---	385
<i>Drosophila melanogaster</i>	SNVQ*RREL-	---	---	---	---	*****	---	455
<i>Drosophila pseudoobscura</i>	SNLQ*RQKL-	---	---	---	---	*****	---	464
<i>Drosophila yakuba</i>	SNAQ*RRAL-	---	---	---	---	*****	---	455
<i>Plasmodium berghei</i>	YFY-NKDN-H	E-Q--	---	---	---	*****	---	400
<i>Plasmodium falciparum</i>	YFYNNQNNLH	E-EFS-	---	---	---	*****	---	395
<i>Plasmodium vivax</i>	FFY-*N-KLH	T-Q--	---	---	---	*****	---	404
<i>Cryptosporidium hominis</i>	***NIKDHLD	--IISA-	---	---	---	*****	---	428
<i>Cryptosporidium parvum</i>	***NIKDHLD	--IISA-	---	---	---	*****	---	428

* * * * *

<i>Bos taurus</i>	NR****FCG*	DKEK*ACEQP	THWRI	*****	*****	***AY*****	*****	429
<i>Rattus norvegicus</i>	---****	---Q---*	---	---	---	---	---	429
<i>Mus musculus</i>	---****	---Q---*	---	---	---	---	---	429
<i>Homo sapiens</i>	---****	---Q---*	---	---	---	---	---	429
<i>Pan troglodytes</i>	---****	---Q---*	---	---	---	---	---	429
<i>Gallus gallus</i>	---****	E---*N---	---	---	---	---	---	438
<i>Xenopus tropicalis</i>	---****	---*N---	---	---	---	---	---	435
<i>Danio rerio</i>	---****	-V-Q*E---	---	---	---	---	---	444
<i>Oryzias latipes</i>	---****	-A-*E---	---	---	---	---	---	421
<i>Takifugu rubripes</i>	---****	-S-*E---	---	---	---	---	---	386
<i>Ciona intestinalis</i>	---****	I-S*	---	---	---	---	---	434
<i>Ciona savignyi</i>	---****	I-A*	---	---	---	---	---	433
<i>Caenorhabditis briggsae</i>	---****	L-P*	---	---	---	---	---	424
<i>Caenorhabditis elegans</i>	---****	L-P*	---	---	---	---	---	424
<i>Anopheles gambiae</i>	-T****	--KS	---	---	---	---	---	451
<i>Bombyx mori</i>	---****	L-P*	---	---	---	---	---	405
<i>Drosophila melanogaster</i>	-T****	--K*	---	---	---	---	---	490
<i>Drosophila pseudoobscura</i>	-T****	I-K*	---	---	---	---	---	495
<i>Drosophila yakuba</i>	-T****	--K*	---	---	---	---	---	490
<i>Plasmodium berghei</i>	*****	L--*	---	---	---	---	---	450
<i>Plasmodium falciparum</i>	*****	L-S*	---	---	---	---	---	447
<i>Plasmodium vivax</i>	*****	L--*	---	---	---	---	---	455
<i>Cryptosporidium hominis</i>	-KCNEV---	---	---	---	---	---	---	461
<i>Cryptosporidium parvum</i>	-KCNEV---	---	---	---	---	---	---	461

*

<i>Bos taurus</i>	*****	*****	*****	429
<i>Rattus norvegicus</i>	*****	*****	*****	429
<i>Mus musculus</i>	*****	*****	*****	429
<i>Homo sapiens</i>	*****	*****	*****	429
<i>Pan troglodytes</i>	*****	*****	*****	429
<i>Gallus gallus</i>	*****	*****	*****	438
<i>Xenopus tropicalis</i>	*****	*****	*****	435
<i>Danio rerio</i>	*****	*****	*****	444
<i>Oryzias latipes</i>	*****	*****	*****	421
<i>Takifugu rubripes</i>	*****	*****	*****	386
<i>Ciona intestinalis</i>	*****	*****	*****	434
<i>Ciona savignyi</i>	*****	*****	*****	433
<i>Caenorhabditis briggsae</i>	*****	*****	*****	424
<i>Caenorhabditis elegans</i>	*****	*****	*****	424
<i>Anopheles gambiae</i>	*****	*****	*****	451
<i>Bombyx mori</i>	*****	*****	*****	405
<i>Drosophila melanogaster</i>	*****	*****	*****	490
<i>Drosophila pseudoobscura</i>	*****	*****	*****	495
<i>Drosophila yakuba</i>	*****	*****	*****	490
<i>Plasmodium berghei</i>	FGFPSHFPSN	LSIYRKKYIP	FA****	472
<i>Plasmodium falciparum</i>	YNFLSHFPNN	RSIYRKEYIT	NT****	469
<i>Plasmodium vivax</i>	YDFAHYYPQN	VSIYRRGERR	EEGRLM	481
<i>Cryptosporidium hominis</i>	*****	*****	*****	461
<i>Cryptosporidium parvum</i>	*****	*****	*****	461

2 Conservation des sites de *N*-glycosylation de Pofut1, relation structure/fonction

2.1	PUBLICATION 2	173
2.2	RESULTATS COMPLEMENTAIRES	175
2.2.1	<i>Expression tissulaire de l'enzyme bovine Pofut1</i>	175
2.2.2	<i>Expression tissulaire de l'enzyme bovine Pofut2</i>	178
2.2.3	<i>Synthèse et conclusion</i>	181

Lors de notre précédente étude, nous avons remarqué que l'enzyme bovine Pofut1 possédait deux sites potentiels de *N*-glycosylation, tout comme l'ensemble des enzymes de Mammifères. Parmi ces deux sites, le premier (N⁶⁵ chez *Bos taurus*) est conservé chez toutes les espèces pour lesquelles nous avons identifié le gène *Pofut1* dans les bases de données, exception faite des nématodes *Caenorhabditis briggsae*, *C. elegans* et *Strongyloides ratti*, de l'Insecte de l'ordre des Hémiptères, *Homalodisca coagulata* et du cnidaire de la classe des Anthozoaires, *Acropora millepora*. Le deuxième site (N¹⁶³ chez *Bos taurus*) est seulement présent dans les séquences de Mammifères et de l'Insecte de l'ordre des Lépidoptères *Bombyx mori*.

Dans le but d'étudier le rôle potentiel de la *N*-glycosylation dans l'activité de l'enzyme Pofut1 bovine, nous avons dans un premier temps démontré que celle-ci est effectivement diglycosylée *in vivo* sur ces deux sites (N⁶⁵ et N¹⁶³), et *in vitro* lorsqu'elle est exprimée par des cellules COS-1 transfectées. Puis, des analyses fonctionnelles des protéines mutées sur l'un, l'autre, ou les deux résidus asparagine ont été réalisées.

L'étude des protéines Pofut1 mutées indique que le second site de *N*-glycosylation est pratiquement facultatif pour l'activité de la protéine. En effet, l'absence de *N*-glycane au second site de *N*-glycosylation ne diminue l'activité *O*-fucosyltransférase *in vitro* que de 26,2%. Ce résultat est cohérent avec la non conservation de ce site chez toutes les espèces de Vertébrés et son absence chez les Invertébrés. Le premier site de *N*-glycosylation est par contre essentiel pour la stabilité et la bonne conformation de la protéine Pofut1 bovine. Sa mutation entraîne son accumulation et son agrégation dans le réticulum endoplasmique.

Ces travaux sont présentés de manière détaillée dans l'article ci-après, publié dans la revue FEBS Journal en 2007.

2.1 Publication 2

**The two *N*-glycans present on bovine Pofut1 are
differently involved in its solubility and activity**

Loriol, C., Audfray, A., Dupuy, F., Germot, A., and Maftah, A.

FEBS Journal, 2007, vol 274, 1202-1211

2.2 Résultats complémentaires

2.2.1 Expression tissulaire de l'enzyme bovine Pofut1

Grâce à nos anticorps polyclonaux anti-Pofut1, nous avons pu établir le profil d'expression de cette enzyme au sein des treize tissus prélevés chez le bovin adulte (Fig.52.). Ce sont ces mêmes tissus qui ont été auparavant utilisés pour la caractérisation du profil d'expression des variants transcriptionnels (cf. Publication 1). De la même façon, le profil d'expression de Pofut1 dans les huit tissus d'embryon bovin a été étudié (Fig.52.). Les résultats, avec d'une part les trois adultes, et d'autre part les trois embryons, étant les mêmes, nous ne présentons que ceux obtenus avec un adulte et un embryon. Une partie de ces résultats est présentée dans la publication 2.

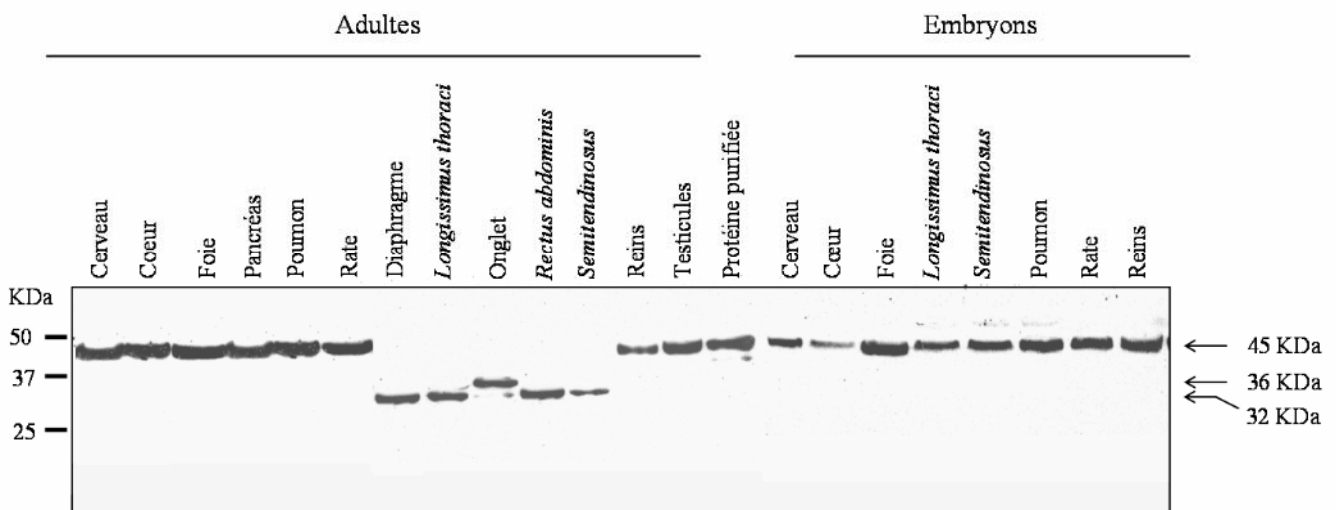


Figure 52 : Profil d'expression tissulaire de la protéine bovine Pofut1.

Treize tissus adultes et huit tissus embryonnaires ont été analysés. 50µg d'extraits protéiques tissulaires sont déposés par piste. 100 ng de la protéine Pofut1 bovine exprimée dans *E. coli* et purifiée servent de témoin. Les tailles des protéines révélées dans les tissus sont indiquées par des flèches sur la droite de la figure.

Comme nous l'avons vu dans le chapitre précédent, la protéine Pofut1 qui a servi d'antigène pour la production des anticorps, porte encore le peptide signal et une étiquette (His)₆ à chacune de ses extrémités. Sa masse moléculaire déduite de la séquence peptidique est de 46,9 KDa. Dans tous les tissus bovins adultes et embryonnaires, à l'exception des muscles squelettiques adultes, nous révélons une bande protéique de taille similaire. La taille prédite de la protéine Pofut1 active et diglycosylée (cf. Résultats complémentaire publication 1), après clivage du peptide signal dans le réticulum endoplasmique est de 45 KDa. On peut par conséquent conclure que tous les tissus, sauf les muscles chez l'adulte, expriment une protéine Pofut1 portant deux *N*-glycanes. Ceci est confirmé par les résultats obtenus dans la publication 2 où la déglycosylation partielle par l'enzyme PNGase F de la protéine recombinante Pofut1 bovine, produite dans les cellules COS-1, génère trois entités protéiques correspondant respectivement à des protéines di-, mono- et déglycosylées. Ce même traitement réalisé pendant 20 heures sur un extrait de protéines totales de foie bovin produit une entité protéique de 41 KDa correspondant à la taille attendue de l'enzyme Pofut1 sans son peptide signal et sans glycanes. Aucune protéine issue de la traduction des variants *Pofut1b* à *1e* (cf. Publication 1) n'a été mise en évidence dans les tissus bovins. *In vivo*, ces transcrits, non traduits, pourraient jouer un rôle dans la régulation du niveau d'expression de l'enzyme Pofut1 active (Lareau *et al.*, 2004). En effet, de plus en plus d'études tendent à montrer que de nombreux ARN non codants (ARNnc) seraient présents dans les cellules eucaryotes. Ceux-ci seraient impliqués dans des processus de régulation de l'expression génique à plusieurs niveaux : ARN interférence, déméthylation de l'ADN, régulation de la structure de la chromatine (Costa, 2005). L'exemple le plus connu est l'inactivation du chromosome X chez les Mammifères par l'ARNnc XIST (Rastan, 1994). Des dérégulations du niveau d'expression de ces ARNnc seraient impliquées dans l'apparition de certaines pathologies cancéreuses (Srikantan *et al.*, 2000) et neurologiques (Runte *et al.*, 2004).

Une protéine d'environ 32 KDa est reconnue dans chacun des extraits protéiques des muscles squelettiques adultes quelque soit leur métabolisme, lent/oxydatif (Diaphragme, Onglet, *Rectus abdominis*) ou rapide/glycolytique (*Semitendinosus*, *Longissimus thoraci*). La bande correspondante est cependant de très faible intensité dans l'onglet où une autre bande protéique d'environ 36 KDa est détectée. Donc, malgré la présence de transcrits *pofut1a* dans les muscles squelettiques chez l'adulte, la protéine correspondante Pofut1A n'est pas révélée. Aucun des autres transcrits caractérisés dans l'étude précédente ne peut expliquer la présence de protéines Pofut1 de cette taille. Afin d'identifier ces protéines, des expériences

d'immunoprécipitation utilisant l'anticorps anti-Pofut1 ont été conduites. Ces expériences n'ont malheureusement pas été concluantes, sans doute à cause du faible niveau d'expression de ces enzymes. Il est certes possible que la présence de ces protéines de faible masse moléculaire soit liée à la traduction d'un ou plusieurs autres variants transcriptionnels non encore identifiés. Néanmoins, cette hypothèse semble peu probable. En effet, étant donné leur faible taille potentielle, nous aurions dû les amplifier lors de nos études par RT-PCR. Il semble peu probable également que ce marquage soit lié à une dégradation protéolytique de la protéine Pofut1 lors de la préparation des échantillons, les extractions protéiques étant réalisées simultanément et en présence de puissants inhibiteurs de protéases. De plus, il est peu vraisemblable que ce type de dégradation génère une entité protéique unique. L'hypothèse la plus vraisemblable est qu'il se produise un clivage protéolytique spécifique au sein des muscles squelettiques adultes. On sait en effet, que certaines protéases sont spécifiques de ces tissus, comme la Calpaïne musculaire (Bartoli et Richard, 2005). Ce clivage conduirait à la présence dans ces tissus d'une forme inactive de l'enzyme Pofut1. Ceci peut être relié avec la forte activité des récepteurs Notch lors de la formation des muscles squelettiques au cours du développement embryonnaire, et à leur activité basale dans les tissus musculaires adultes (Conboy *et al.*, 2003). Pour conforter cette hypothèse, il serait nécessaire d'effectuer des tests d'activité enzymatique à partir d'extraits protéiques totaux de muscles embryonnaires et adultes.

L'étude de l'expression tissulaire de Pofut1 révèle donc que la forme active est invariablement présente dans certains tissus bovins (cerveau, cœur, foie, poumon, rate et reins) de l'adulte et de l'embryon. A l'inverse, les tissus musculaires squelettiques adultes n'expriment pas les mêmes formes de la protéine. Il serait intéressant d'analyser l'expression de la protéine Pofut1, dans des tissus musculaires squelettiques provenant d'embryons et de nouveaux-nés à des stades de développement différents, ceci afin de mettre en évidence la transition entre l'expression de la forme active et de la forme tronquée.

Dans le but de révéler si des formes atypiques de la protéine Pofut1 existaient dans les muscles squelettiques d'organismes appartenant à d'autres espèces, nous avons utilisé l'anticorps anti-Pofut1 pour une réaction croisée sur des protéines extraites de tissus de deux autres espèces de Mammifères. Les tissus adultes examinés sont le cœur et le *Semitendinosus* de souris, et deux muscles squelettiques de porc (muscles de la cuisse et longe) (Fig.53.).

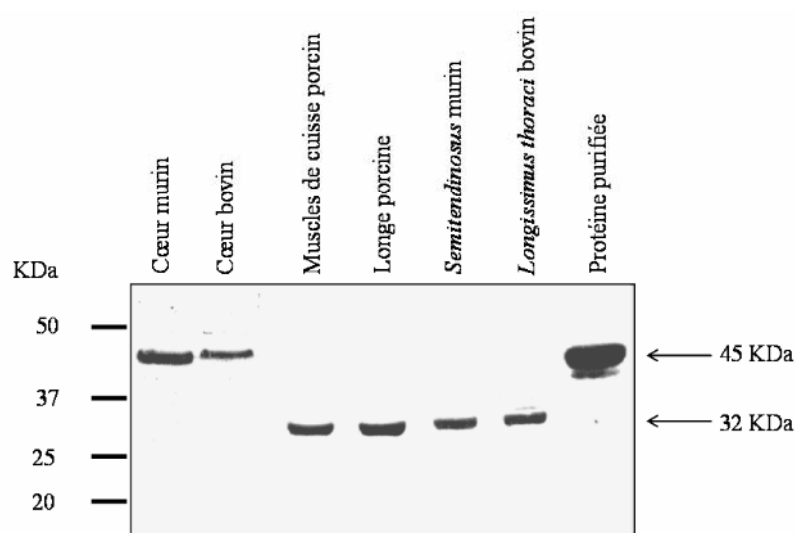


Figure 53 : Expression de Pofut1 dans des tissus de Mammifères.
L'analyse a porté sur 50 μ g d'extrait protéique total d'un tissu et sur 100ng de protéine purifiée.

Une protéine qui correspondrait à Pofut1, est reconnue par l'anticorps dans l'extrait protéique de cœur murin. Sa masse est similaire à celle de Pofut1A détectée dans un extrait de cœur bovin. Les extraits de muscles squelettiques de souris et de porc présentent une forme protéique de taille comparable à celle détectée dans les muscles squelettiques bovins adultes. Ainsi, les muscles squelettiques de ces deux autres espèces n'exprimeraient pas non plus la forme active de l'enzyme. Cette analyse confirme donc la présence chez d'autres espèces de Mammifères d'une forme tronquée de la protéine Pofut1. Il existerait donc un mécanisme conservé d'inactivation de la protéine Pofut1 dans les muscles squelettiques de Mammifères adultes. La caractérisation de cette forme protéique revêt donc un intérêt crucial, d'une part pour mieux comprendre la fonction biologique de la forme active, et d'autre part pour mettre en évidence si cette forme tronquée possède elle-même une fonction biologique indépendante de l'activité *O*-fucosyltransférase.

2.2.2 Expression tissulaire de l'enzyme bovine Pofut2

Cette étude a été conduite en parallèle avec celle portant sur l'expression de Pofut1. En effet, grâce aux anticorps polyclonaux anti-Pofut2, nous avons étudié l'expression tissulaire de la protéine bovine Pofut2, au sein des treize tissus adultes et des huit tissus embryonnaires bovins cités plus haut (Fig.54.).

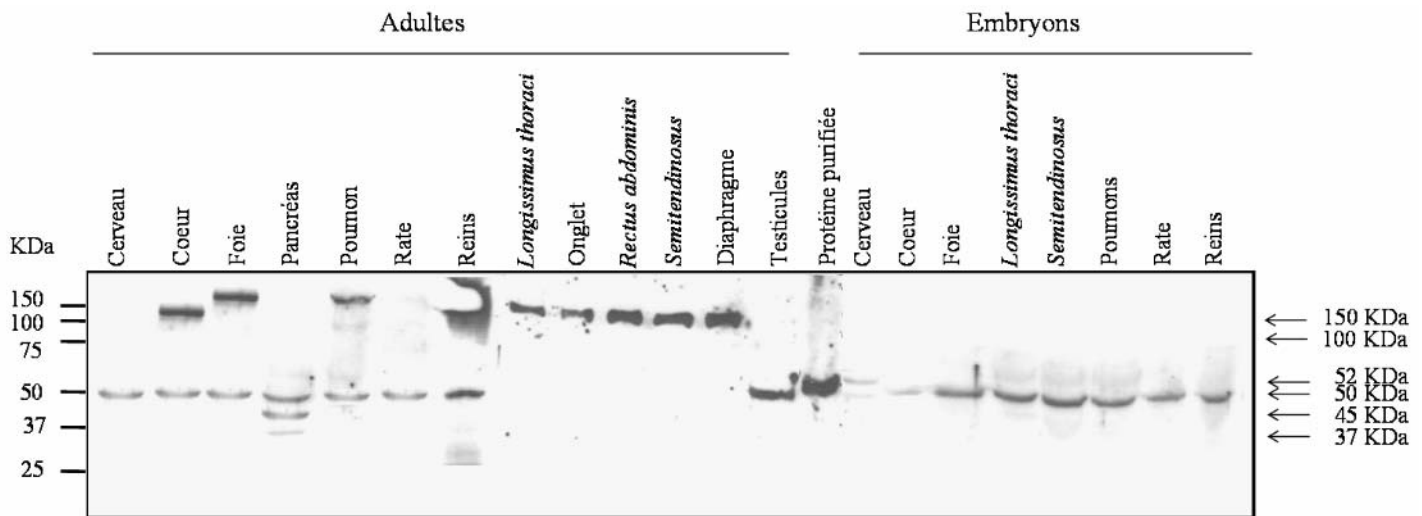


Figure 54 : Profil d'expression tissulaire de la protéine bovine Pofut2.

Treize tissus adultes et huit tissus embryonnaires ont été analysés. 50µg d'extraits protéiques tissulaires sont déposés par piste. 100 ng de la protéine Pofut2 bovine exprimée dans *E. coli* et purifiée servent de témoin. Les tailles des protéines révélées sont indiquées par des flèches sur la droite de la figure.

Comme l'illustre la figure 54, le profil d'expression de Pofut2 dans les divers tissus analysés est plus complexe et encore plus inattendu que celui de Pofut1. La protéine Pofut2 qui a servi d'antigène pour la production des anticorps correspond à la forme Pofut2B et porte encore le peptide signal et une étiquette (His)₆ à chacune de ses extrémités. Sa masse moléculaire déduite de la séquence peptidique est de 52 KDa. Dans tous les tissus, sauf les muscles squelettiques chez l'adulte, une protéine d'environ 50 KDa est détectée. Celle-ci pourrait correspondre à une forme triglycosylée de l'enzyme active Pofut2A ou diglycosylée de l'enzyme Pofut2B. Cependant, cette dernière hypothèse semble peu probable étant donné l'activité non significative de la forme Pofut2B *in vitro* (cf. Résultats complémentaire publication 1). Dans le cerveau embryonnaire, une forme protéique supplémentaire est détectée à une taille légèrement supérieure (~52 KDa); elle pourrait correspondre à la forme triglycosylée de Pofut2B. Dans le pancréas adulte, il existe deux formes supplémentaires de 45 et 37 KDa. Celles-ci ne peuvent pas correspondre aux produits des variants transcriptionnels *Pofut2c*, *d* ou *e*, dont la traduction conduirait à la même protéine non glycosylée de 18 KDa. Ces formes protéiques peuvent donc résulter de la traduction d'un autre variant transcriptionnel, qui serait spécifique de ce tissu. D'autres protéines de très haut poids moléculaires sont détectées dans le foie, les poumons et les reins adultes (~150 KDa) et dans le cerveau et tous les muscles squelettiques adultes (~100 KDa). Ces protéines sont d'ailleurs les seules à être reconnues par l'anticorps dans les muscles squelettiques chez

l'adulte, alors que les variants transcriptionnels du gène *Pofut2a* y sont présents. Ces résultats ont été confirmés par des expériences indépendantes qui révèlent toujours le profil présenté dans la figure 54. Ces formes de haut poids moléculaires pourraient correspondre à des dimères (100 KDa) et des trimères (150 KDa) d'enzymes Pofut2 ou être dues à une association de Pofut2 avec d'autres protéines présentes dans les tissus adultes, association qui serait résistante aux conditions dénaturantes de l'électrophorèse. Ces résultats mettent en évidence un mode de fonctionnement différent de la *O*-fucosyltransférase 2 dans les tissus embryonnaires et adultes. Des expériences d'immunoprécipitation seront toutefois nécessaires pour confirmer la nature des formes protéiques de haut poids moléculaires présentes dans les tissus musculaires squelettiques adultes.

Dans le but de révéler si des formes atypiques de la protéine Pofut2 existaient dans les muscles squelettiques d'organismes appartenant à une autre espèce nous avons utilisé l'anticorps pour une réaction croisée sur des protéines extraites de deux tissus murins, du poumon et du *Semitendinosus* (Fig.55.).

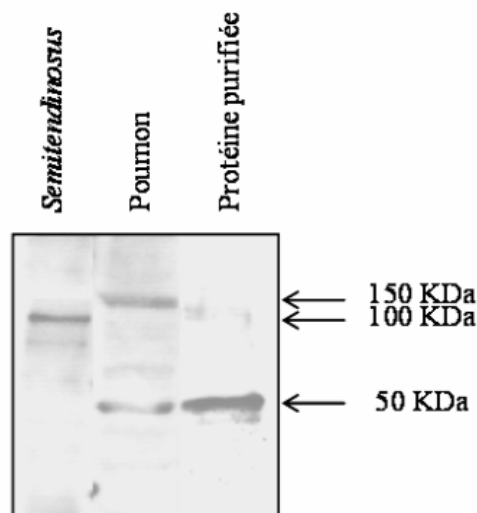


Figure 55 : Expression de Pofut2 dans deux tissus murins. L'analyse a porté sur 50µg d'extrait protéique total d'un tissu et sur 100ng de protéine purifiée.

Une protéine qui correspondrait à Pofut2 est reconnue par l'anticorps, dans l'extrait de poumon murin, à une taille de 50 KDa. Comme dans le cas du bovin, ce tissu présente une autre forme protéique à un poids moléculaire supérieur (~150 KDa). Les muscles squelettiques de souris n'exprimeraient pas non plus la forme active de l'enzyme puisqu'une protéine de 100 KDa est majoritairement détectée. Ce résultat semble confirmer l'existence

d'un mode d'action conservé de la *O*-fucosyltransférase 2 dans les muscles squelettiques adultes bovins et murins.

2.2.3 Synthèse et conclusion

Durant cette étude, nous avons donc déterminé le profil d'expression des enzymes Pofut1 et Pofut2 bovines dans treize tissus d'adultes et huit tissus d'embryons. Mis à part les formes atypiques détectées dans les muscles squelettiques adultes, seules les protéines Pofut1 et Pofut2 actives, issue de la traduction des variants *1a* et *2a*, seraient détectées *in vivo* sous une forme respectivement di- et triglycosylée. Les variants *Pofut1b* à *1e* et *Pofut2b* à *2e* mis en évidence lors de notre précédente étude ne sont donc pas traduits *in vivo* (sauf peut-être Pofut2B dans le cerveau adulte) et pourraient jouer un rôle au niveau transcriptionnel dans la régulation de l'expression des enzymes Pofut1 et Pofut2 actives.

L'étude de l'impact des sites de *N*-glycosylation de l'enzyme Pofut1 bovine révèle que la mutation du second site de *N*-glycosylation ne modifie pas la solubilité de l'enzyme mais diminue son activité catalytique. La mutation du premier site de *N*-glycosylation entraîne quant à elle la disparition de la protéine de la phase soluble d'un extrait protéique de cellule COS-1 transfectées, suggérant son agrégation dans le réticulum endoplasmique. En effet, seul un traitement de la fraction insoluble de ces extraits par de l'urée 10M permet sa libération et sa détection en western blot. Ce traitement n'autorisant pas les tests d'activité enzymatique, il nous a été impossible de vérifier si l'enzyme produite possédait encore une activité *O*-fucosyltransférase.

Une étude similaire à celle réalisée sur Pofut1, portant sur les fonctions des *N*-glycanes de l'enzyme Pofut2 est envisagée. Etant donné l'importance démontrée de la *O*-fucosylation des domaines EGF dans les processus développementaux et dans la myogenèse, mes efforts se sont d'avantage centrés sur Pofut1.

3 Pofut1 : un rôle dans la différenciation de cellules musculaires en culture ?

3.1	MISE EN PLACE DE LA CULTURE PRIMAIRE DE CELLULES SATELLITES BOVINES ET DES VECTEURS D'EXPRESSION EUCARYOTE	187
3.1.1	<i>Isolement, culture et différenciation.....</i>	187
3.1.2	<i>Construction des vecteurs eucaryotes pour les transfections</i>	189
3.1.3	<i>Transfection des cellules bovines en culture primaire.....</i>	194
3.2	ETUDE DU ROLE DE POFUT1 DANS LA DIFFERENCIATION DE CELLULES C2C12.	199
3.2.1	<i>Construction du vecteur de surexpression</i>	199
3.2.2	<i>Culture et transfection.....</i>	200
3.2.3	<i>Premiers résultats obtenus</i>	205
3.3	MATERIELS ET METHODES	215

Les cellules satellites sont considérées comme les cellules souches du muscle squelettique adulte. Localisées à la périphérie des fibres, elles ont la propriété de remodeler le muscle squelettique adulte en participant à sa croissance et/ou sa réparation, puisque capables de générer des cellules qui vont entrer en différenciation et des cellules qui vont adopter un phénotype immature. Comme nous l'avons vu dans l'exposé bibliographique, le récepteur Notch est un récepteur transmembranaire dont l'activation induit l'expression de gènes impliqués dans la détermination cellulaire et tissulaire ainsi que dans le maintien de l'état immature des cellules souches au cours du développement de nombreux tissus. Il a été montré que la surexpression de formes constitutivement actives du récepteur Notch1 dans des myoblastes prolifératifs (cellules satellites activées), inhibe leurs différenciations (Nofziger *et al.*, 1999). De plus, la perte progressive de la capacité régénératrice des cellules satellites de souris, au cours du vieillissement, est en partie due à une diminution de la signalisation induite par le récepteur Notch (Conboy *et al.*, 2003). Afin de mieux caractériser le rôle du récepteur Notch, dans la détermination et l'entrée en prolifération des cellules satellites, Kitzmann, Cornac et leurs collaborateurs ont caractérisé l'expression de Notch1 dans des myoblastes prolifératifs de la lignée myoblastique murine C2C12 (communication au Club Muscle et Développement III, 16-17 Décembre 2004, Paris). Dans cette étude, Notch1 s'est révélé être responsable du maintien d'une sous population de cellules satellites dans un état immature. L'inhibition de son activité (par l'utilisation d'inhibiteurs de γ -secrétases) entraîne une fusion de ces cellules avec les myotubes préexistants, conduisant ainsi à une hypertrophie musculaire.

La *O*-fucosylation des membres de la famille Notch est indispensable à leur activité. En effet, les *O*-fucosylglycanes présents à la fois sur les récepteurs Notch et sur leurs ligands vont moduler leurs interactions et vont donc participer au réglage des voies de signalisation intra cytoplasmique conduisant, dans la cellule myoblastique, à l'expression de gènes inhibiteurs de la différenciation. C'est pourquoi, nous avons pensé que la modulation de l'expression du gène *Pofut1*, responsable de la *O*-fucosylation de Notch et ses ligands, aurait un impact sur la différenciation des cellules musculaires en culture.

Pour vérifier cette hypothèse, nous avons entrepris d'étudier les répercussions d'une modulation de l'expression du gène *Pofut1* (surexpression ou sous-expression) sur le déroulement du programme myogénique de cellules myoblastiques en culture, notamment sur l'expression des facteurs myogéniques de la famille bHLH. Pour ce faire, il a fallu développer

au sein du laboratoire les techniques d'isolement, de culture et de différenciation de cellules satellites bovines et construire les différents vecteurs d'expression eucaryote nécessaires aux transfections transitoires. En même temps, nous avons développé la même approche sur les cellules myoblastiques de la lignée murine C2C12. Ce dernier modèle cellulaire est également utilisé au sein de l'équipe pour étudier l'expression du glycogénome au cours de la différenciation des myoblastes en myotubes.

3.1 Mise en place de la culture primaire de cellules satellites bovines et des vecteurs d'expression eucaryote

Le but ultime de cette thèse étant d'évaluer le rôle joué par l'enzyme bovine Pofut1 au cours de la différenciation de cellules myoblastiques en culture, nous avons mis en place l'isolement et la culture primaire de cellules satellites extraites du muscle squelettique *Semitendinosus* d'embryons bovins.

3.1.1 Isolement, culture et différenciation

Les cellules satellites bovines, une fois extraites du muscle squelettique (cf. annexe III qui constitue la partie « Matériels et Méthodes » de cette troisième partie), sont placées dans un milieu de culture qui les engage dans le cycle cellulaire pour devenir des myoblastes. Quarante huit heures après l'isolement, les cellules satellites musculaires sont triées de façon à éliminer les fibroblastes isolés de façon simultanée, lesquels, de par leur vitesse de prolifération prendraient rapidement le dessus sur les cellules musculaires. Ce tri s'effectue au FACS (Fluorescence Activated Cell Sorter) grâce à un marquage utilisant un anticorps anti-CD56 couplé à la phycoérythrine (PE) (Becton-Dickinson, Le Pont de Claix, France). L'anticorps reconnaît de façon spécifique les myoblastes, qui portent à leur surface le marqueur CD56 ou PSA-NCAM (polysialic acid-neural cell adhesion molecule) intervenant dans les processus de migration neuronale et musculaire, alors que les fibroblastes en sont dépourvus. Comme contrôle isotypique de fluorométrie, nous utilisons des IgG1 de souris couplées à l'isothiocyanate de fluorescéine (FITC) (Becton-Dickinson). Ils permettent d'estimer et de corriger la liaison non spécifique des anticorps aux cellules (Fig.56.).

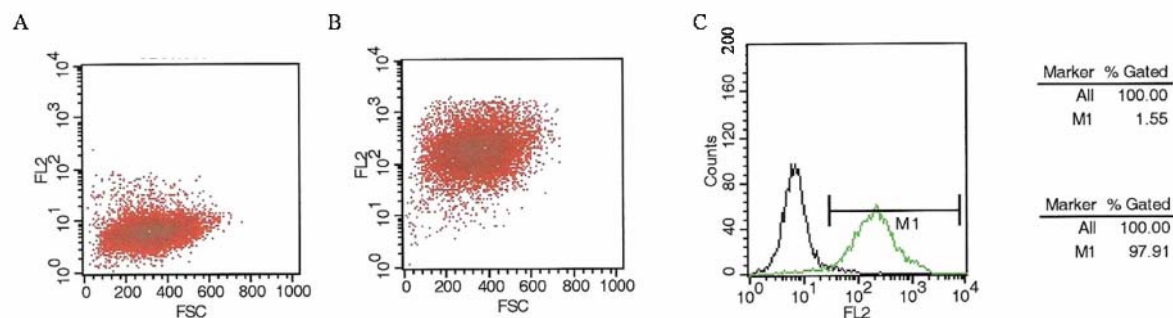


Figure 56 : Résultat d'un tri sur un ensemble de cellules isolées d'un muscle squelettique d'embryon bovin.

A, Cytogramme biparamétrique, rapport taille vs fluorescence verte des cellules marquées avec l'anticorps isotypique couplé FITC; B, Cytogramme biparamétrique, rapport taille vs fluorescence orange des cellules marquées avec l'anticorps anti-CD56 couplé PE; C, Histogramme monoparamétrique, rapport fluorescence vs nombre de cellules. 1,55% des cellules sont marquées avec l'anticorps isotypique et 97,91% avec l'anticorps anti-CD56. Ainsi, 96,36% des cellules sont des myoblastes.

L'isolement de cellules satellites bovines a également été effectué à partir d'un muscle bovin adulte. Le nombre de cellules collectées, après tri, étant beaucoup plus faible (environ 15% contre 95% en moyenne avec un muscle d'embryon), nous avons décidé de poursuivre notre étude sur des cellules provenant d'embryons.

Les cellules satellites, triées de façon stérile, peuvent être remises en culture. Ces cellules myoblastiques primaires en culture, comme les cellules de lignées, ont la capacité de se différencier *in vitro*, c'est-à-dire de fusionner pour former des myotubes. Pour contrôler les différentes étapes de leur différenciation, les cellules sont cultivées dans un premier temps dans un milieu dit de prolifération, contenant 10% de sérum de veaux fœtal (FCS), cette forte concentration de sérum inhibant la différenciation. Lorsque les cellules sont presque à confluence (recouvrant entre 80 et 90% de la flasque de culture), le milieu est remplacé par un milieu contenant 2% de sérum de cheval qui va induire la fusion.

Des immunomarquages avec un anticorps dirigé contre la laminine, fortement exprimée dans les myotubes, permettent de vérifier que les cellules isolées et cultivées sont capables de fusionner pour former des myotubes (Fig.57.).

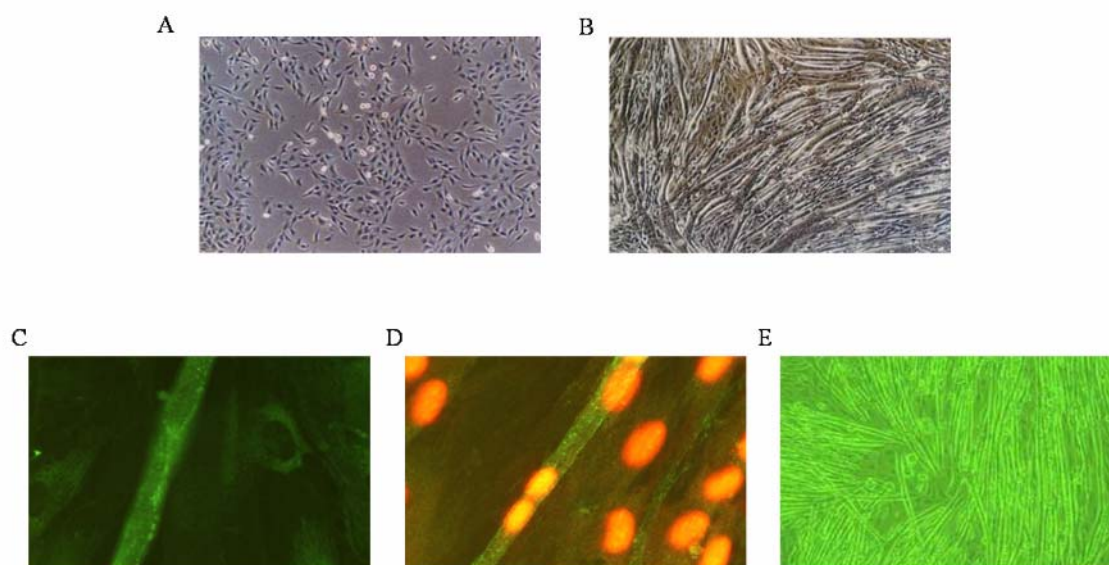


Figure 57 : Photographies en lumière blanche et immunomarquages des myoblastes et myotubes bovins.

A, myoblastes bovins en phase proliférative (X10); B, myotubes formés après 120 heures en milieu de différenciation (X10); C, marquage d'un myotube avec l'anticorps anti-laminine (X40); D, marquage à l'iodure de propidium (IP) afin de visualiser les noyaux (X40); E, réseau de myotubes bovins marqués avec l'anticorps anti-laminine (X10). L'anticorps anti-laminine, produit chez la souris est utilisé comme anticorps primaire. L'anticorps secondaire est un anticorps anti-IgG de souris couplé à la FITC.

En quelques mois, nous avons donc mis en place une culture de myoblastes bovins et nous avons su en maîtriser la culture jusqu'à son terme, la différenciation en myotubes. Néanmoins, comme nous le verrons plus loin, la culture primaire de cellules présente plusieurs inconvénients majeurs.

3.1.2 Construction des vecteurs eucaryotes pour les transfections

Un vecteur d'expression eucaryote permettant de surexprimer l'enzyme Pofut1A bovine, et un autre vecteur dit de shRNA (short hairpin RNA) permettant d'inhiber l'expression d'ARNm du gène *Pofut1a*, ont été construits.

3.1.2.1 Vecteur de surexpression

Conscients du faible taux de transfection généralement atteint avec des cellules en cultures primaires, nous avons choisi d'utiliser un vecteur d'expression eucaryote portant la séquence codante de l'eGFP (ou d'un de ses dérivés). Ce vecteur nous permet de vérifier le

taux de transfection et de trier en cytométrie de flux, les cellules effectivement transfectées. A la suite du tri, les cellules transfectées sont remises en culture et leur différenciation induite 24 heures plus tard. Onze flasques sontensemencées afin de suivre la cinétique de différenciation. Les répercussions de la surexpression de *Pofut1* sur la différenciation d'une population homogène de cellules peuvent alors être observée.

Le choix du vecteur d'expression n'a pas été simple. En effet, fusionner la protéine Pofut1A recombinante directement au marqueur fluorescent (que se soit en N ou C-terminal) entraîne une perte totale de l'activité *O*-fucosyltransférase.

Notre choix s'est donc porté sur le vecteur pVITRO2-IGFP (Cayla-InvivoGen, Toulouse, France) (Fig.58).

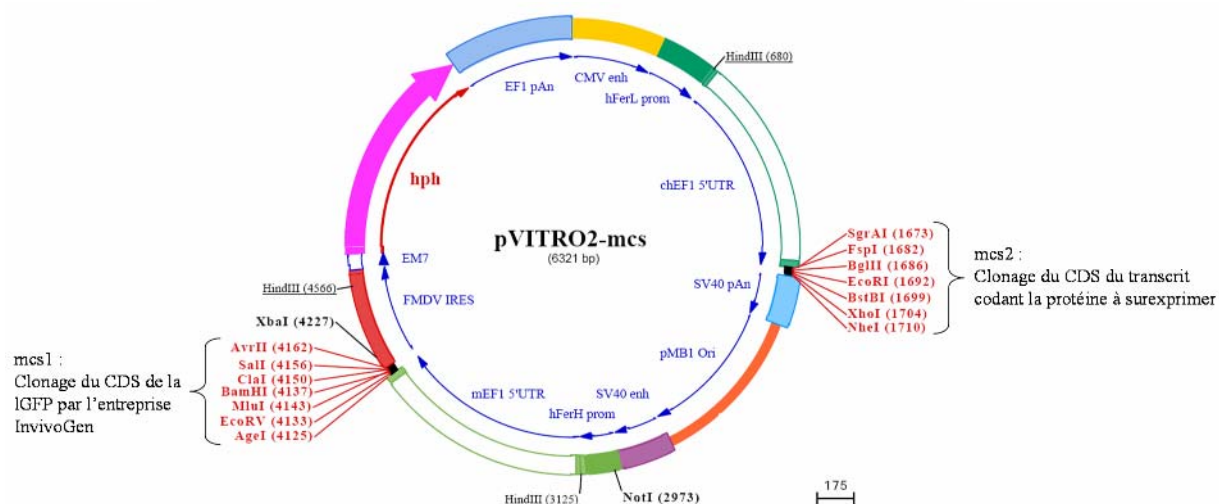


Figure 58 : Vecteur choisi pour la surexpression de l'enzyme Pofut1A dans les cellules primaires de muscle bovin.

Ce vecteur possède deux sites multiples de clonage (mcs). Au mcs1, nous avons demandé à l'entreprise InvivoGen de cloner le CDS du transcript codant la IGFP (un dérivé de la GFP) excitable à 488nm. Au mcs2, nous avons cloné le CDS du transcript bovin *Pofut1a* codant la forme active de la *O*-fucosyltransférase 1.

Ce vecteur possède deux promoteurs composites nommés hFer/m-chEF1. Ils résultent de l'association entre le promoteur ubiquitaire de la ferritine humaine et la région 5'UTR du gène *EF1* de la souris (promoteur avant le mcs1) ou du chimpanzé (promoteur avant le mcs2). Ainsi, l'activation du promoteur est constitutive et non régulée par le taux de fer. Ils sont de plus renforcés respectivement par l'enhancer du promoteur du SV40 ou du CMV. Ces deux promoteurs constitutifs vont permettre l'expression au sein des cellules eucaryotes transfectées, de deux protéines distinctes, d'une part la IGFP dont les longueurs d'onde

d'excitation et d'émission sont les mêmes que celles de la eGFP (soit respectivement 488nm et 510nm), et d'autre part notre protéine d'intérêt Pofut1A.

La séquence codante de *Pofut1a* a été introduite dans le vecteur d'expression eucaryote pVITRO2-IGFP, par clonage directionnel au niveau des sites de restriction *Bg*III-*Xho*I. Le clonage de cette séquence bovine a nécessité, au préalable, son amplification avec un couple d'amorces dont l'une apporte le site de restriction enzymatique *Bg*III, permettant de conserver le cadre de lecture dans ce vecteur et la séquence Kozak de *Pofut1a* bovin (6 paires de bases en amont de l'ATG qui constitue un site de reconnaissance des ribosomes; Kozac, 1989). Ces amorces sont *Pofut1-Bg*III (5'-AGATCTCCCGCTATGGGCGCCGCC-3') et *Pofut1-CDS*3' (5'-TCAGAATTCATCCCGCAGCTGAG-3'). Comme matrice, nous avons pris la construction pcDNA/TOPO3.1-*Pofut1a* utilisée pour la transfection des cellules COS-1 et la réalisation des tests d'activité enzymatique présentées dans la partie 2 des résultats. L'amplification ainsi obtenue a été de nouveau clonée dans le vecteur pcDNA/TOPO3.1. La séquence d'intérêt a été isolée grâce à la double digestion par les enzymes de restriction *Bg*III (site apporté par notre amorce) et *Xho*I (présent sur le vecteur en 3' de l'insert). Elle a été ensuite ligaturée aux mêmes sites dans le vecteur pVITRO2-IGFP. Les constructions ont servi à la transformation des bactéries *Escherichia coli* TOP10F'. Après leur lyse alcaline, les plasmides ont été purifiés par Maxi préparation d'ADN (plasmid maxi Kit, QIAGEN, Les Ulis, France). L'insert a été séquencé dans son intégralité grâce à deux amorces complémentaires du vecteur (pVITRO5' 5'-AGGCCAGCTTGGCACTTGATGTAATTCTC-3'; pVITRO3' 5'-AATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAA-3').

3.1.2.2 Vecteur de shRNA

Le principe de l'interférence par ARN est de fournir à la cellule un petit ARN double brin ayant la même séquence qu'une portion de l'ARN à détruire, en l'occurrence celui de *Pofut1a* bovin. Nous avons opté pour la technique du shRNA qui permet d'éteindre l'expression d'un gène d'intérêt de façon plus stable et plus durable qu'en fournissant directement aux cellules des petits ARN interférents (Fig.59.).

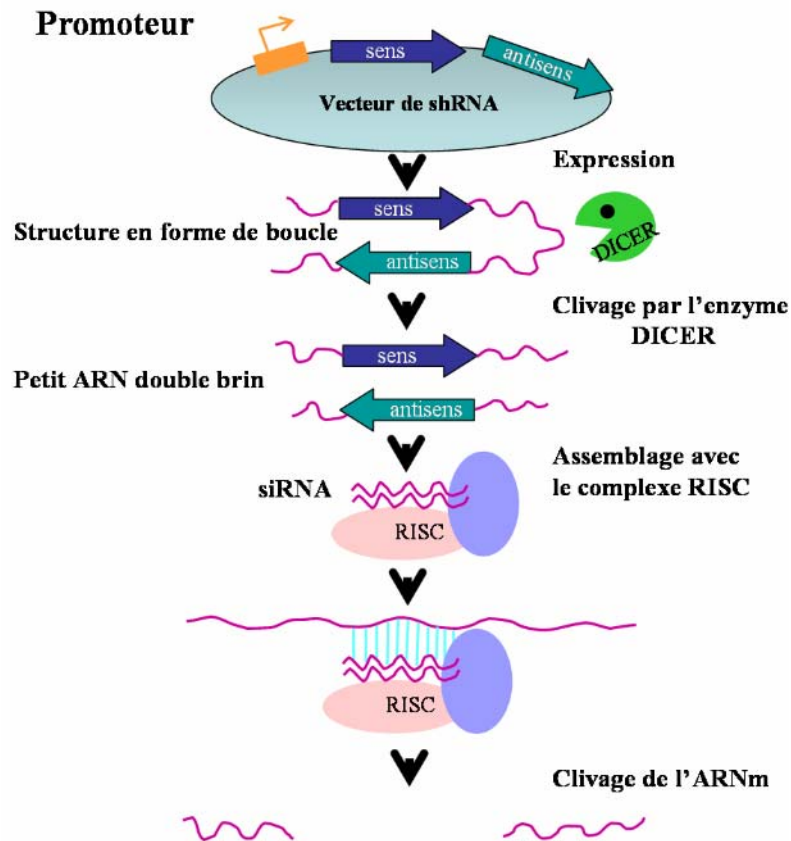


Figure 59 : Principe du shRNA.

Les molécules d'ARN double brin, produites par la transcription du petit insert contenu dans le vecteur, sont d'abord découpées par une RNase, l'éminceuse (Dicer), qui produit des fragments de 19 à 21 nucléotides, les petits ARN interférents. Ceux-ci sont incorporés sous forme simple brin dans un complexe ribonucléoprotéique dénommé RISC (*RNA-induced silencing complex*) où ils servent de guide pour la reconnaissance de la cible. Un appariement parfait avec l'ARNm cible détermine une activité de nucléase spécifique du complexe avec la coupure endonucléotidique et la dégradation des fragments d'ARNm cible.

Un vecteur de shRNA nous a été donné par le Professeur Pamela Stanley (Albert Einstein College of Medicine, New York, USA). Ce vecteur contient un fragment de séquence correspondant à une région de l'exon 3 du gène *Pofut1* murin, en sens et en antisens (Fig.60.).

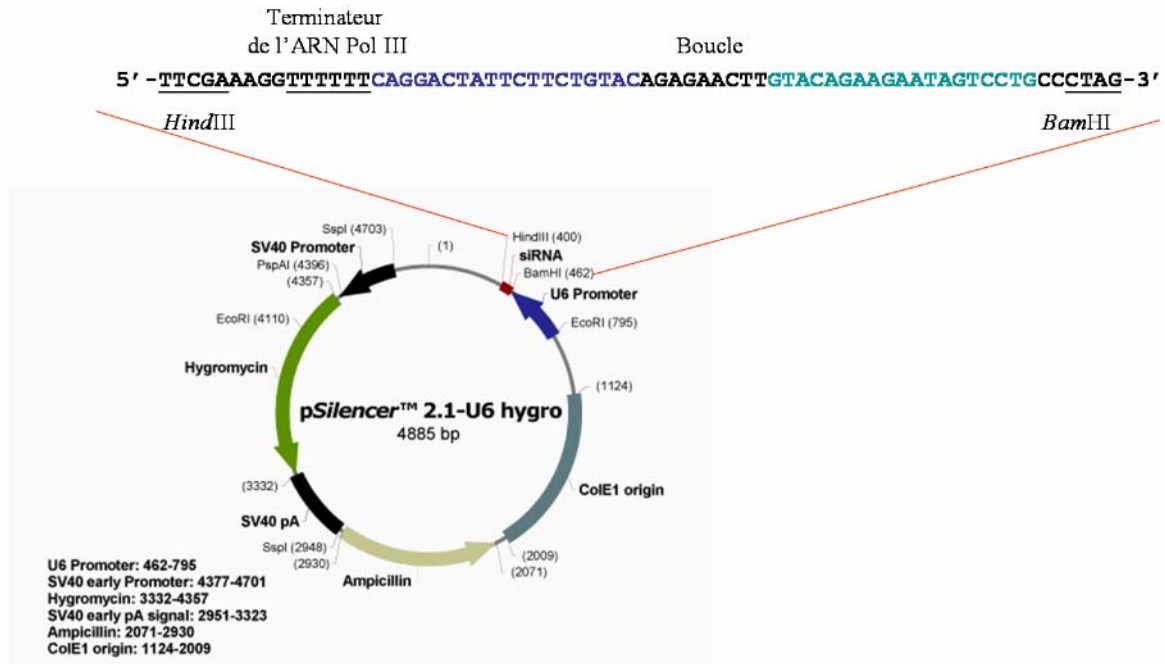


Figure 60 : Vecteur pSilencer 2.1-U6 de shRNA.

Le vecteur contient un fragment (19 pb) de séquence en sens (bleu) et en antisens (vert), homologue à une région de l'exon 3 du gène *Pofut1* murin. L'expression de l'insert est sous le contrôle du promoteur U6.

L'identité parfaite de l'insert avec une portion de la séquence bovine et son efficacité dans l'extinction de l'expression d'ARN *Pofut1* dans les cellules CHO (communication personnelle de P. Stanley) nous ont motivé pour l'utiliser. Néanmoins, ce vecteur ne possédant pas de marqueur permettant de contrôler le taux de transfection et de trier les cellules, nous avons entrepris le clivage de l'insert par les enzymes de restrictions *BamHI* et *HindIII*, de façon à l'insérer au niveau des sites de restrictions *BglIII* (isoschizomère de *BamHI*) et *HindIII* du vecteur pSuper-RNAi (OligoEngine, Seattle, USA) (Fig.61.). Etant donné la faible taille de l'insert, nous n'avons pas réussi à l'isoler même par l'intermédiaire d'une migration sur un gel d'acrylamide. Nous avons donc reconstruit l'insert en utilisant deux amorces complémentaires avec chacune une extrémité débordante, correspondant au site de restriction désiré (siRNA1 : 5'-GATCCCGTCCTGATAAGAAGACATGTTCAAGAGACATGTCTTCTTATCAGGACTTTTTTGGAA-3'; siRNA2 : 5'-AGCTTTTCCAAAAAGTCCTGATAAGAAGACATGTCTCTTG AACATGTCTTCTTATCAGGACGA-3'). Celles-ci ont été hybridées 10min à 96°C, puis le mélange a ensuite été ligaturé dans le vecteur pSuper-RNAi précédemment clivé par les enzymes de restriction *BglIII* et *HindIII*. La vérification de la présence de l'insert a été réalisée

après séquençage grâce à deux amorces complémentaires du vecteur (T7 : 5'-AATACGACTCACTATAG-3' et M13 : 5'-GTAAAACGACGGCCAGT-3').

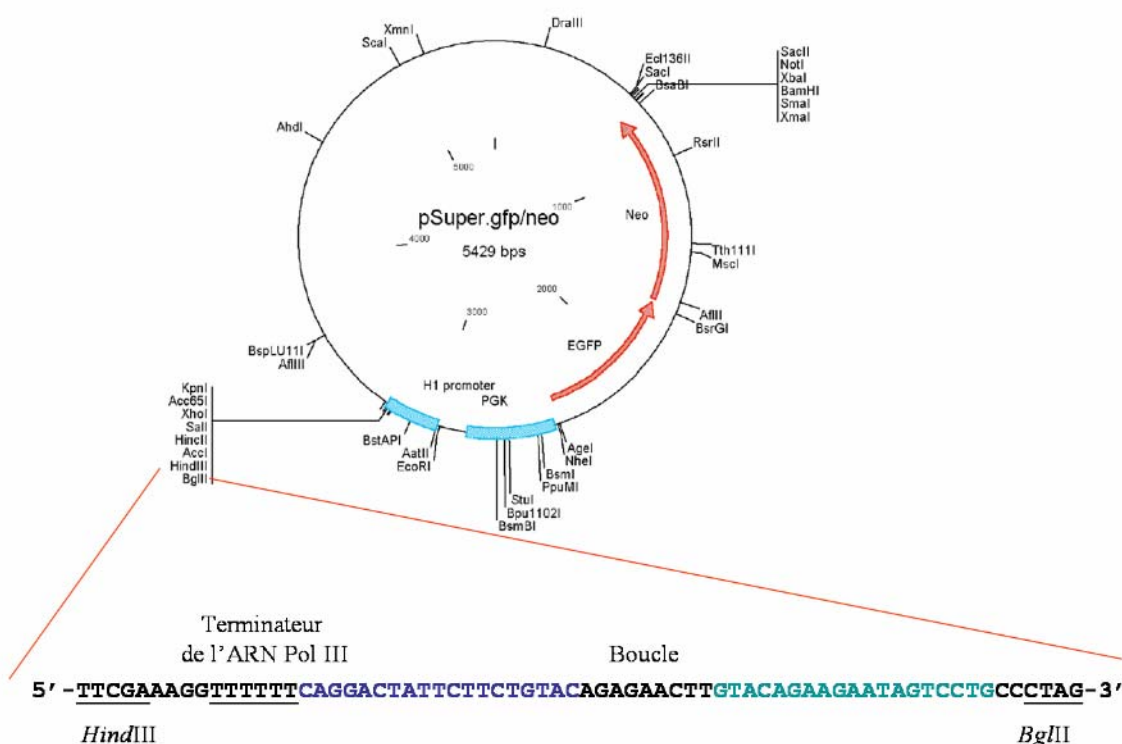


Figure 61 : Vecteur pSuper.gfp de shRNA.

L'expression du shRNA est sous le contrôle du promoteur H1 et celle de l'eGFP sous le contrôle du promoteur fort de la Phosphoglycérate kinase (PGK).

3.1.3 Transfection des cellules bovines en culture primaire

La culture primaire de cellules présente comme intérêt de mieux refléter ce qui se passe *in vivo* par rapport aux lignées cellulaires qui sont modifiées, de façon spontanée ou induite, pour devenir immortelles. Plus le nombre de passages (décollement, dilution et repiquage) augmente et plus les caractéristiques de ces cellules s'éloignent de celles de cellules primaires en culture. Néanmoins, la culture primaire de cellules présente des inconvénients majeurs. Le nombre de passages des cellules est limité, nous ne pouvons repiquer nos cellules myoblastiques bovines que 5 ou 6 fois. Passées ces étapes, les cellules cessent de proliférer et meurent. Ainsi, à chaque fois que nous avons besoin de cellules, il faut en isoler et en cultiver de nouvelles. Le second inconvénient des cultures primaires auquel nous nous sommes heurté est la difficulté de transfection.

Nous avons souhaité commencer notre étude par la surexpression de Pofut1A avec le vecteur pVITRO2-IGFP/*Pofut1a* bovin. Avant toute analyse, il est nécessaire d'avoir suffisamment de cellules transfectées, de façon à pouvoir les trier grâce à l'expression de la IGFP, et les remettre en culture. Pour effectuer notre cinétique de différenciation, il nous faut 11 boîtes de cellules transfectées (flasques de 25cm² ou boîtes de 100mm de diamètre pour que la différenciation se passe normalement et soit observable) qui seront mises en milieu de différenciation une fois les 80% de confluence atteints. Les cellules seront donc récupérées à 11 temps différents recouvrant toutes les étapes de la différenciation de 0% à 100% de myotubes.

Nous avons essayé plusieurs techniques de transfection, notamment grâce à l'utilisation de polymères cationiques comme le jetPEI (polyéthylèneimine, Polyplus-Transfection, Illkirch, France) ou le PolyFect (PolyFect Transfection Reagent, QIAGEN, Les Ulis, France). Seules les transfections effectuées avec des liposomes (lipofection, Fugene6, Roche diagnostics, Meylan, France) nous ont permis d'obtenir un pourcentage de cellules transfectées entre 3 et 4% (Fig.62A.), taux plus faible que celui auquel nous nous attendions. De plus, un premier essai de repiquage de ces cellules transfectées et triées montre qu'au cours de leur prolifération, elles perdent progressivement l'expression de la IGFP (Fig.62B et C.).

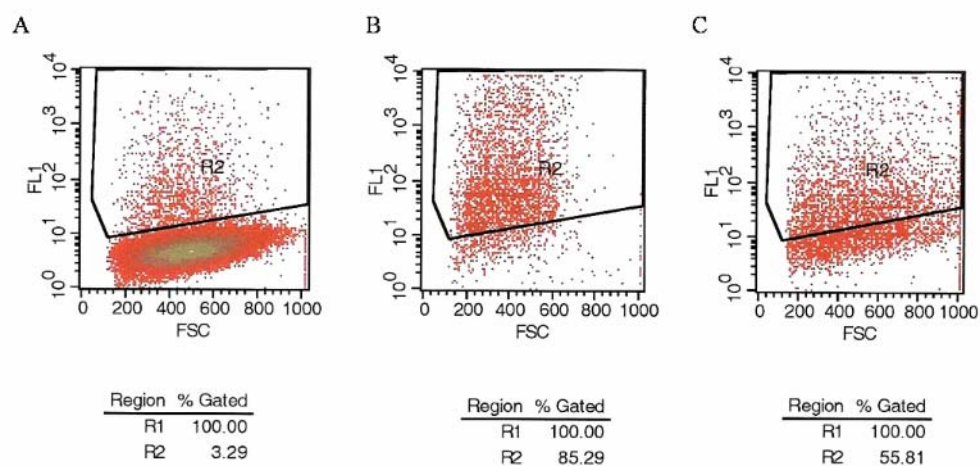


Figure 62 : Cytogrammes biparamétriques (rapport taille vs fluorescence) obtenus lors d'une expérience de tri cellulaire après transfection de culture primaire de myoblastes bovins. A, tri cellulaire 24 heures après transfection avec le plasmide pVITRO2-IGFP/*Pofut1a* bovin. Seulement 3,29% des cellules expriment la IGFP. Après tri stérile, ces cellules sont remises en culture dans deux boîtes de 60mm de diamètre; B, passage au FACS des cellules d'une des deux boîtes 24 heures après tri. 85,29% de cellules expriment toujours la IGFP; C, passage au FACS des cellules de la seconde boîte 48 heures après le tri. Il n'y a plus que 55,81% de cellules qui expriment la IGFP. R1 : pourcentage de cellules prises en compte, R2 : pourcentage de cellules exprimant la IGFP.

Des expériences de western blot sur des extraits protéiques de cellules triées, réalisées avec l'anticorps anti-Pofut1, nous ont permis de constater que la protéine recombinante était produite dans les cellules primaires mais que son expression, comme celle de la IGFP, chutait rapidement (Fig.63).

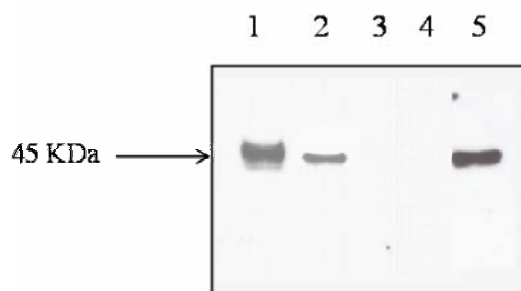


Figure 63 : Western blot anti-Pofut1 réalisé sur 50µg d'extraits protéiques de cellules primaires transfectées par le plasmide pVITRO2-IGFP-*Pofut1a* bovin et triées.

1, expression de Pofut1 au moment du tri; 2, expression de Pofut1, 24 heures après tri; 3, expression de Pofut1, 72 heures après tri; 4, expression de Pofut1 dans les cellules non transfectées; 5, 100ng de Pofut1A purifiée.

Etant donné la perte rapide d'expression de la protéine Pofut1A, il faudrait pouvoir induire la différenciation des cellules dès leur remise en culture après tri, pour que celles-ci n'aient plus le temps de se diviser et de perdre leur plasmide. Pour ce faire et afin de réaliser une cinétique de différenciation, il faudrait pour chacune de nos 11 boîtes de culture, repiquer entre 1.10^6 et 2.10^6 cellules transfectées. Ceci est simplement inenvisageable avec un pourcentage de cellule transfectées généralement inférieur à 5%.

Pour vérifier si le faible taux de transfection n'était pas inhérent à notre vecteur, nous avons effectué des contrôles de transfection avec le vecteur pcDNA/TOPO3.1-LacZ (Invitrogen, Carlsbad, USA). Il permet une réaction chromogène des cellules transfectées grâce à l'utilisation de X-Gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside), substrat de la β -galactosidase. Comme vous pouvez le constater sur les photographies de la figure 64, le taux de transfection est aussi, dans ce cas, extrêmement faible.

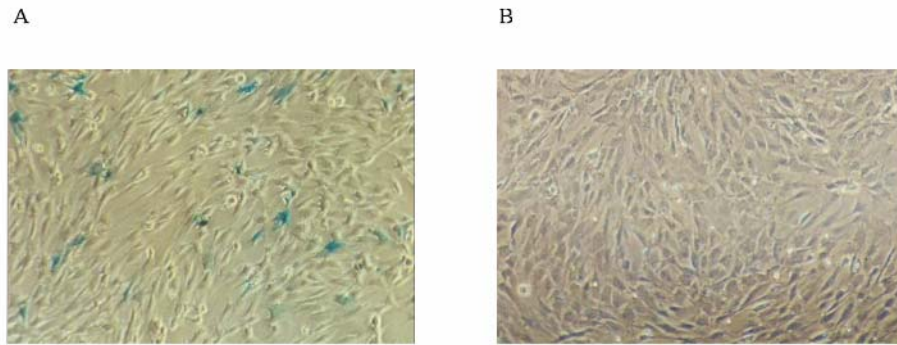


Figure 64 : Photographies en lumière blanche de cellules myoblastiques bovines en phase proliférative.

A, cellules transfectées avec le pcDNA/TOPO3.1-LacZ et traitées au Xgal; les cellules exprimant la β -galactosidase deviennent bleues (X10). B, cellules non transfectées (X10).

Au vu des difficultés rencontrées, nous n'avons pas tenté, pour l'instant, de transfecter ces cellules avec le vecteur de shRNA. Pour obtenir une première information sur l'effet de la modulation de la *O*-fucosylation lors de la myogenèse, nous avons donc décidé de travailler avec une lignée cellulaire, les cellules myoblastiques murines C2C12.

3.2 Etude du rôle de *Pofut1* dans la différenciation de cellules C2C12

Dans le but d'appréhender le rôle joué par l'enzyme *Pofut1* au cours de la différenciation de cellules myoblastiques murines de la lignée C2C12, il nous a fallu, dans un premier temps, construire les outils moléculaires nécessaires aux transfections. En l'occurrence, un vecteur de surexpression contenant le CDS de *Pofut1* murin. Le vecteur de shRNA précédemment construit (Fig.61.) contenant une séquence homologue aux séquences bovine et murine pourra nous être utile dans cette étude.

3.2.1 Construction du vecteur de surexpression

Grâce à la sélection d'amorces spécifiques de part et d'autre du CDS de *Pofut1* murin, nous avons pu isoler, par PCR, l'ADNc codant la forme active de l'enzyme en utilisant comme matrice des ADNc de cœur obtenus après extraction et rétro-transcription d'ARN totaux. Le gène *Pofut1* étant suffisamment exprimé chez la souris, une PCR primaire suffit. Afin de cloner l'ADN amplifié dans le vecteur pVITRO2-IGFP, nous avons utilisé une amorce en 5' possédant le site de restriction de l'enzyme *Bgl*III (souligné) et la séquence Kozak de *Pofut1* murin (*Pofut1*-Mus-*Bgl*III : 5'-AGATCTGTTCGACATGGGCGCCGCC-3'), et une amorce en 3' complémentaire à la région terminale du CDS, comportant le codon stop (*Pofut1*-Mus-CDS3' : 5'-TCAAAATTCATCCCGAAGCTGGGAGG-3'). L'amplification obtenue a tout d'abord été clonée dans le pcDNA/TOPO3.1. La double digestion de ce vecteur par *Bgl*III et *Xho*I autorise ensuite la ligature de l'insert dans le vecteur pVITRO2-IGFP linéarisé par ces mêmes enzymes.

3.2.2 Culture et transfection

Les techniques de culture et de différenciation des cellules C2C12 sont depuis longtemps maîtrisées au sein du laboratoire (Voir annexe III).

Pour transférer ces cellules, nous avons choisi d'utiliser la technologie Amaxa (Amaxa, Cologne, Allemagne) qui offre le premier système efficace de transfert non-viral de gènes dans des lignées cellulaires difficiles à transférer. Elle utilise le principe reconnu de l'électroporation, cependant modifié et amélioré, puisque l'ADN plasmidique est transféré directement dans le noyau des cellules, indépendamment de la division cellulaire, on parle donc de nucléoporation. Elle combine certains paramètres électriques spécifiques, fournis par l'appareil « Nucleofector », et des solutions de transfection spécifiques à chaque type de cellule.

Deux jours après transfection, on procède au tri des cellules en cytométrie de flux. Le but de la manipulation, tout comme avec les cellules primaires bovines, est de pouvoir récupérer un nombre suffisant de cellules transfectées.

3.2.2.1 Difficultés rencontrées

Même avec l'utilisation du système Amaxa, nous avons rencontré de nombreuses difficultés pour la transfection des cellules C2C12.

Avec le vecteur pVITRO2-IGFP/*Pofut1*-murin nous n'avons pu obtenir un pourcentage de transfection suffisant, puisqu'il ne dépassait pas les 10% (Fig.65.).

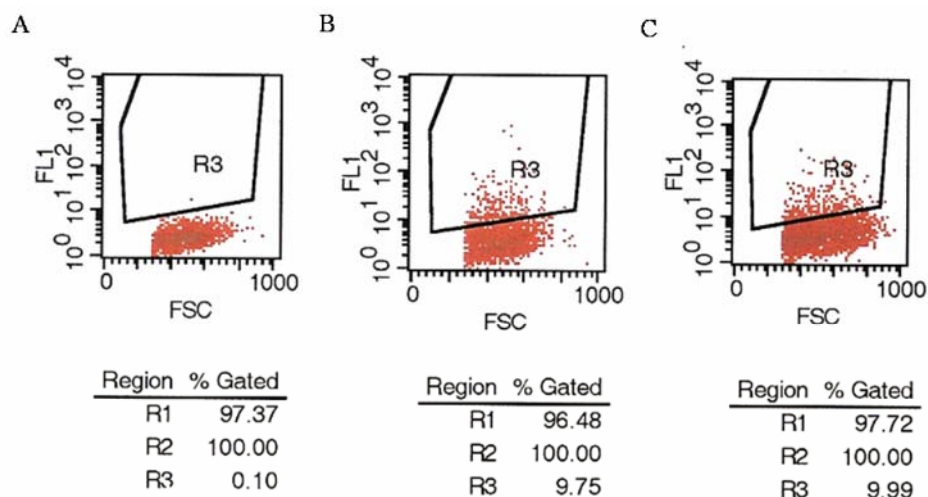


Figure 65 : Cytogrammes biparamétriques (rapport taille vs fluorescence) obtenus lors d’une expérience de tri cellulaire après transfection des cellules C2C12 avec le plasmide pVITRO2-IGFP/*Pofut1* murin.

A, cellules non transfectées; B, cellules transfectées avec le plasmide sans insert. 9,75% des cellules expriment la IGFP; C, cellules transfectées avec le plasmide recombinant. 9,99% des cellules vivantes expriment la IGFP. R1 : pourcentage de cellules vivantes, R2 : pourcentage de cellules vivantes prises en compte, R3 : pourcentage de cellules vivantes exprimant la IGFP.

Le vecteur pVITRO2-IGFP est long de plus de 7Kb, ce qui peut expliquer le faible niveau de transfection. En effet, les transfections réalisées dans les mêmes conditions avec le vecteur témoin fourni par Amaxa, pmaxGFP (Fig.66.) qui fait seulement 3,5Kb, atteignent un pourcentage de l’ordre de 60% (Fig.67.).

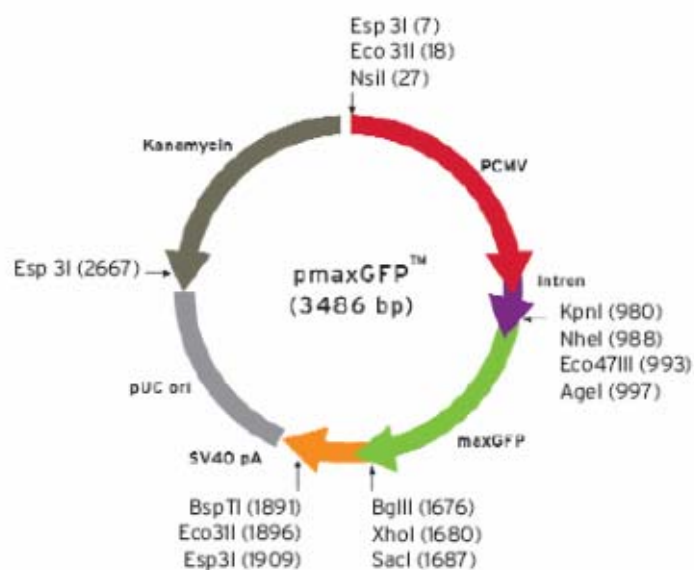


Figure 66 : Vecteur témoin pmaxGFP fourni avec le kit de transfection Amaxa. Ce vecteur est relativement petit : 3486 pb. L'expression de l'eGFP est sous le contrôle du promoteur fort CMV.

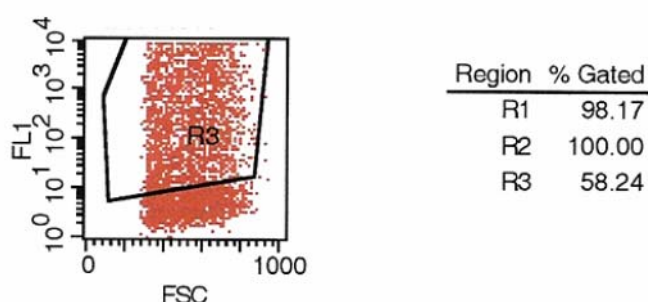


Figure 67 : Cytogramme biparamétrique (rapport taille vs fluorescence) obtenus lors d'une expérience de tri cellulaire après transfection des cellules C2C12 avec le plasmide pmaxGFP. 58,24% des cellules expriment l'eGFP. R1 : pourcentage de cellules vivantes, R2 : pourcentage de cellules prises en compte, R3 : pourcentage de cellules exprimant l'eGFP.

Dans le vecteur pmax, l'eGFP est sous le contrôle du promoteur fort du gène précoce du cytomégalo virus humain (pCMV) permettant l'expression de protéines recombinantes dans toutes les cellules animales, et non du promoteur hFerH/m-EF1 comme dans le pVITRO2-IGFP. Ce dernier est pourtant un promoteur fort qui permet également l'expression de gènes dans toutes les cellules animales (communication InvivoGen, Toulouse, France). Il est cependant possible qu'il ne soit pas un promoteur optimal pour les cellules C2C12.

Les expériences de transfection avec le vecteur de shRNA (pSuper-RNAi de 5,4 Kb) ne nous ont pas permis d'obtenir un pourcentage suffisant de cellules transfectées, puisqu'il est seulement d'environ 7%. Ce vecteur n'ayant pas une taille trop importante, il est possible que le faible taux de cellules exprimant l'eGFP résulte de la présence du promoteur PGK (Phosphoglycérate kinase). Tout comme le CMV, il s'agit d'un promoteur ubiquitaire fort; il est cependant envisageable que celui-ci ne soit pas un promoteur optimal pour les C2C12. Pour vérifier cette hypothèse, des expériences de remplacement de ce promoteur par celui du CMV sont actuellement en cours au laboratoire.

3.2.2.2 Utilisation d'autres vecteurs

Etant donné les difficultés rencontrées avec l'utilisation de vecteurs permettant l'expression de l'eGFP ou la IGFP, et donc le tri des cellules, nous avons décidé de continuer notre étude avec les vecteurs pcDNA/TOPO3.1 pour la surexpression et pSilencer (Fig.60.) pour la sous-expression.

Dans le vecteur pcDNA/TOPO3.1, l'expression de la protéine *Pofut1* est sous le contrôle du promoteur CMV. L'expression du shRNA dans le pSilencer est sous le contrôle du promoteur U6. Ces deux promoteurs forts permettent l'expression transitoire dans de très nombreux types cellulaires.

Ces vecteurs ne contenant pas de CDS codant une protéine fluorescente, il nous est impossible de déterminer le taux de transfection cellulaire. Cependant, nous avons vérifié, dans les deux cas, le niveau d'expression des transcrits de *Pofut1 murin* par RT-PCR semi-quantitative en temps réel (Fig.68.), grâce à l'utilisation d'une sonde TaqMan (Applied Biosystems, Norwalk, USA) spécifique du transcrit.

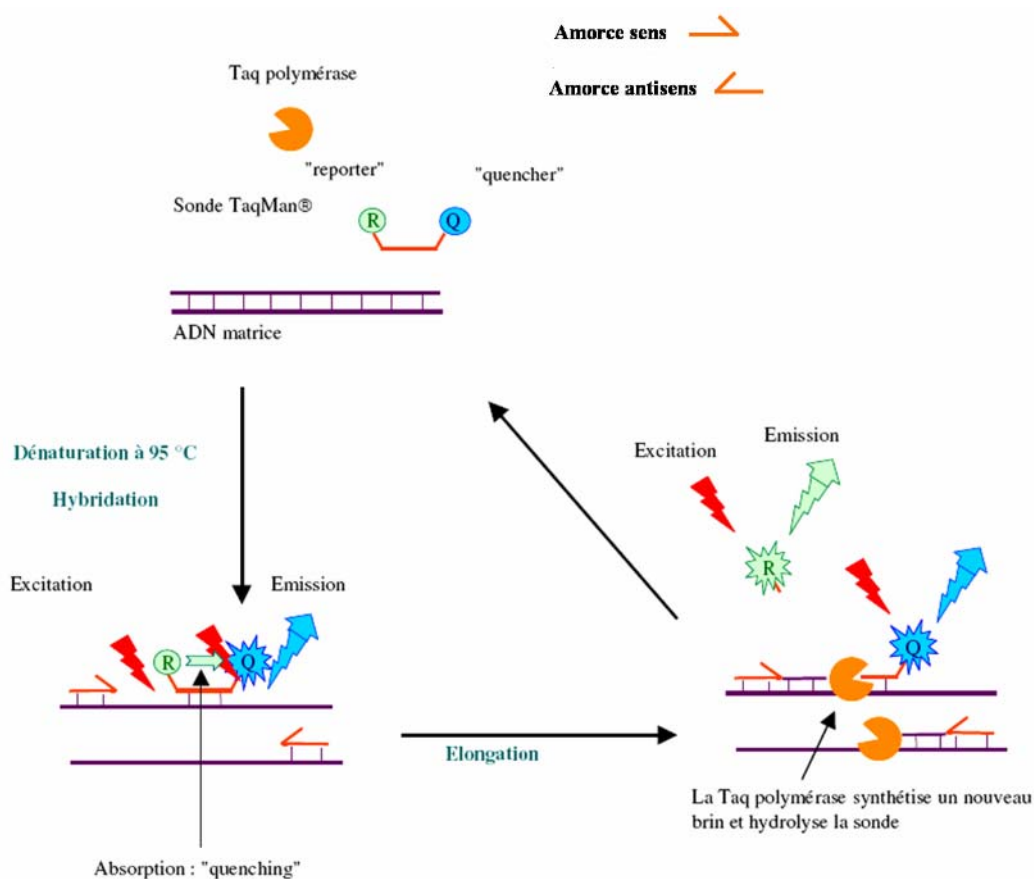


Figure 68 : Principe de la PCR semi-quantitative en temps réel avec une sonde TaqMan.

La méthode de détection à partir de la sonde TaqMan est basée sur l'utilisation de l'activité exonucléasique 5'-3' de l'enzyme *Taq* polymérase. Les sondes sont marquées par un « Reporter » qui émet de la fluorescence uniquement s'il est suffisamment loin d'un « Quencher ». En effet, le Quencher a la propriété d'absorber l'énergie émise par le Reporter, par un mécanisme appelé FRET (*Fluorescence Resonance Energy Transfert*). Lorsque l'enzyme synthétise un nouveau brin, la sonde est hydrolysée; le reporter se retrouve alors en solution et il émet dans sa longueur d'onde caractéristique. Cette technique présente l'avantage d'être plus spécifique par rapport à une PCR classique, de par la présence de la sonde supplémentaire, et de permettre la détection de plusieurs cibles simultanément en utilisant des Reporters qui émettent à des longueurs d'onde différentes.

Etant donné la forte homologie entre les enzymes Pofut1 bovine et murine, des expériences de western blot avec l'anticorps anti-Pofut1, nous permettront également de vérifier l'expression de la protéine Pofut1 dans les cellules C2C12 transfectées et non transfectées.

3.2.3 Premiers résultats obtenus

3.2.3.1 Surexpression de *Pofut1* dans les cellules C2C12

3.2.3.1.1 Transfection et induction de la différenciation

Dans le but d'appréhender le rôle de l'enzyme *Pofut1* dans la différenciation des cellules C2C12, 11 millions de cellules ont été nucléoporées avec la construction pcDNA/TOPO3.1-*Pofut1* murin, puis remises en culture dans 11 flasques de culture cellulaire de 25cm², à raison de 1 million de cellules par flasque. En parallèle, des cellules non transfectées, sont remises en culture de la même façon. La mise en milieu de différenciation s'effectue 24 heures après la transfection. Au même moment, des cellules transfectées et des cellules non transfectées sont décollées et congelées, elles représentent le temps zéro de l'expérience noté T0. De même, des cellules transfectées et des cellules non transfectées sont collectées et congelées à différents temps après la mise en milieu de différenciation. Les différents points, T6, T12, T18, T24, T48, T72, T96, T120, T192 et T264 heures, ont été choisis de façon à couvrir les différentes étapes de la différenciation de 0% à 100% de myotubes.

3.2.3.1.2 Vérification de l'expression de *Pofut1*

Dans un premier temps, les ARN totaux sont extraits des cellules C2C12 (RNeasy Mini Kit, QIAGEN, Les Ulis, France). La qualité et la concentration des ARN sont vérifiées grâce au system Agilent, technologie dont le principe est la microphorèse capillaire sur puce. Dix microgrammes de ces ARN sont ensuite rétro-transcrits (High capacity cDNA Archive Kit, Applied Biosystems, Norwalk, USA).

L'expression des transcrits *Pofut1* est vérifiée par RT-PCR semi-quantitative en temps réel. L'analyse est effectuée, en triplicat, sur les 11 points des deux cinétiques (Fig.69).

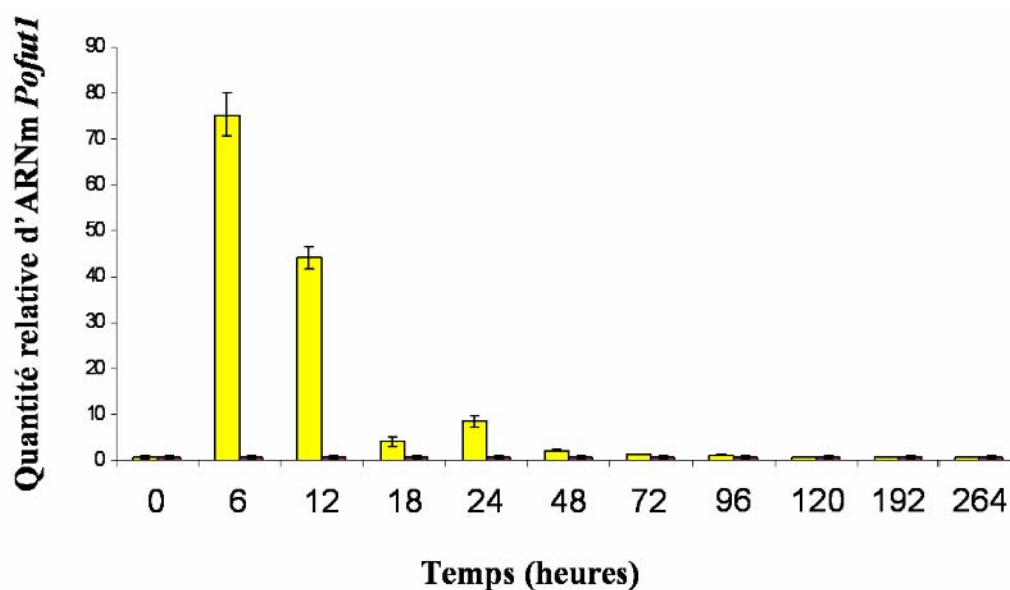


Figure 69 : Analyse en RT-PCR semi-quantitative des transcrits de *Pofut1* dans les cellules C2C12 non transfectées (marron) et dans les cellules transfectées avec la construction pcDNA/TOPO3.1-*Pofut1* murin (jaune).

Le gène *Pofut1* est faiblement exprimé dans les cellules C2C12 quel que soit leur état de différenciation. En effet, les premières amplifications ne sont détectées qu'au terme d'environ 31 cycles de PCR (Ct). Le logiciel analysant les résultats des RT-PCR quantitatives considère qu'un gène n'est pas exprimé, ou très faiblement, lorsque des amplifications ne sont détectées qu'au terme de 33 Ct. Dans les cellules transfectées avec la construction pcDNA/TOPO3.1-*Pofut1* murin, en fixant comme référence zéro le niveau d'expression détectée dans les cellules non transfectées, on remarque une forte expression des transcrits *Pofut1* dans les heures suivant la transfection (T0 à T24). A T6, le facteur de surexpression est d'environ 80X. Comme nous avons choisi d'effectuer des transfections transitoires, le plasmide utilisé ne s'intègre donc pas dans leur génome de la cellule. Il se dégrade donc rapidement et se dilue à mesure que les cellules se divisent. L'expression des transcrits *Pofut1* décroît donc par la suite pour revenir à son niveau basal à T72.

Une étude par western blot avec l'anticorps anti-Pofut1, sur des extraits protéiques totaux de cellules C2C12 correspondant à tous les points de la cinétique (transfectées et non transfectées), est actuellement en cours au laboratoire. Une analyse préliminaire effectuée sur des lysats cellulaires, 30 heures après transfection (T6), confirme l'expression de la protéine Pofut1 recombinante par rapport à un extrait protéique de cellules non transfectées (Fig.70.).

La protéine endogène des cellules C2C12 n'est pas détectée, ce qui reste cohérent avec le faible taux de transcrits *Pofut1* détectés dans ces cellules.

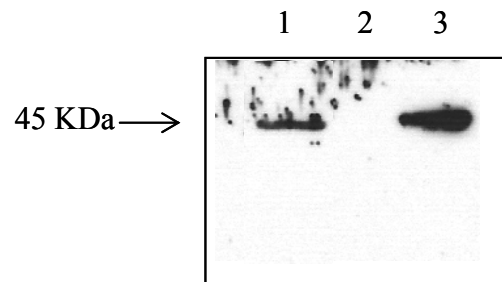


Figure 70 : Western blot réalisé avec l'anticorps anti-Pofut1 sur 50µg d'extraits de cellules C2C12, 24 heures après transfection avec le plasmide pcDNA/TOPO3.1-*Pofut1* murin. 1, expression de Pofut1 dans les cellules transfectées; 2, expression de Pofut1 dans les cellules non transfectées; 3, 100ng de protéine recombinante purifiée utilisée comme témoin positif.

La surexpression du gène *Pofut1* au cours d'une telle différenciation cellulaire ne permet pas d'observer de différences phénotypiques significatives entre les cellules des deux cinétiques de différenciation (Fig.71 et 72.). Toutefois, il semblerait que le nombre de myotubes présents à partir de T48 soit moins important lorsque le gène *Pofut1* est surexprimé. Le nombre exact et la taille des myotubes obtenus n'ont pas encore été quantifiés précisément.

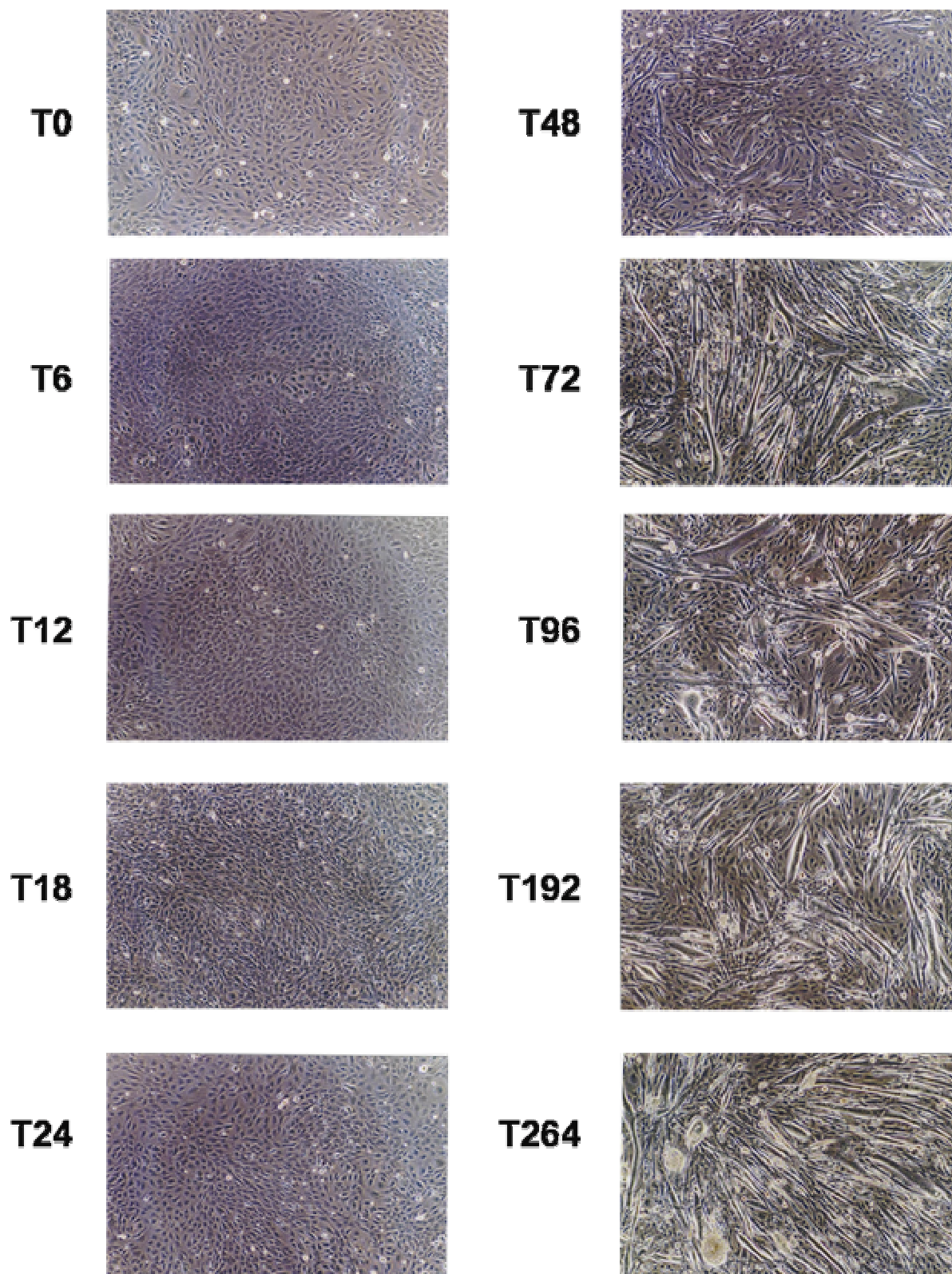


Figure 71 : Photographies, au microscope inversé (10X) en lumière blanche, des cellules C2C12 non transfectées, au cours de la cinétique de différenciation.

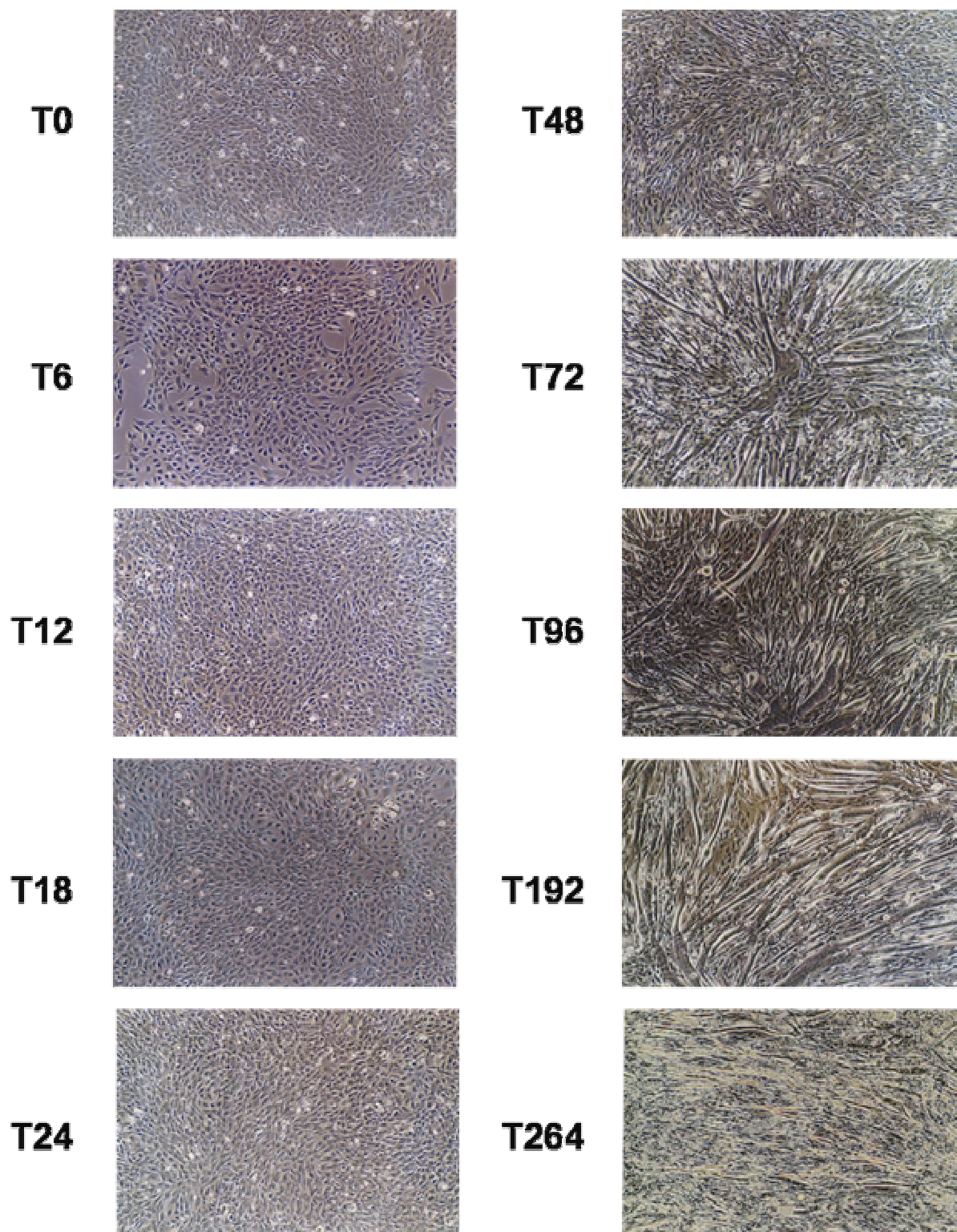


Figure 72 : Photographies, au microscope inversé (10X) en lumière blanche, des cellules C2C12 transfectées par le vecteur pcDNA/TOPO3.1-*Pofut1* murin, au cours de la cinétique de différenciation.

3.2.3.1.3 Répercussion de la surexpression de *Pofut1* sur l'expression des facteurs myogéniques bHLH

Pour appréhender à l'échelle moléculaire la répercussion de la surexpression du gène *Pofut1* sur la différenciation des cellules C2C12, une étude portant sur le niveau d'expression des transcrits du gène *CD34* et de quatre marqueurs myogéniques à domaine bHLH (*MyoD*, *Myf5*, *MRF4* et *Myogénine*) a été réalisée (Fig.73.). Ces manipulations ont été effectuées par Mathilde Janot (étudiante en Master II recherche) et Fabrice Dupuy (Maître de Conférences).

La molécule CD34 est un marqueur des cellules indifférenciées. Chez l'Homme, par exemple, il est le marqueur le plus couramment utilisé pour caractériser les cellules souches et les progéniteurs. Il est exprimé, entre autres, dans les cellules souches hématopoïétiques adultes et dans des cellules satellites du muscle squelettique adulte (Beauchamp *et al.*, 2000). Dans la cinétique de différenciation des cellules C2C12 non transfectées, la quantité de transcrits *CD34* diminue à partir de T24. Cette diminution coïncide avec l'apparition des premiers myotubes (T48). En effet, la diminution du nombre de cellules indifférenciées et l'augmentation de la taille des cellules myogéniques après induction de la différenciation, s'accompagnent d'une diminution de la quantité de l'antigène CD34 (Tamaki *et al.*, 2003). Dans les cellules transfectées avec la construction pcDNA/TOPO3.1-*Pofut1* murin, le marqueur CD34 affiche un profil d'expression différent. En effet, on constate une augmentation du niveau d'expression des transcrits jusqu'à T48 puis, globalement, une diminution jusqu'à la fin de la cinétique (T48 à T264). Ce profil peut-être mis en relation avec la proportion de cellules qui vont rester dans un état indifférencié. En effet, il est envisageable que la surexpression de *Pofut1* empêche les cellules transfectées de se différencier en myotubes, de ce fait une plus grande proportion de cellules continue à exprimer le marqueur CD34 par rapport aux cellules au même stade mais n'ayant pas subi la transfection. Par la suite, la diminution du niveau d'expression de *Pofut1*, due aux divisions cellulaires et à la destruction du plasmide, autorise les cellules à rentrer dans le processus de différenciation, d'où la diminution de l'expression du CD34. Ces résultats suggèrent qu'une *O*-fucosylation intense dans les cellules C2C12 entraîne un blocage des cellules dans un état indifférencié et provoque ainsi un retard dans le processus de différenciation. Ceci pourrait être mis en relation avec une hyper-activation du récepteur Notch, impliqué dans le maintien d'une population cellulaire dans un état quiescent.

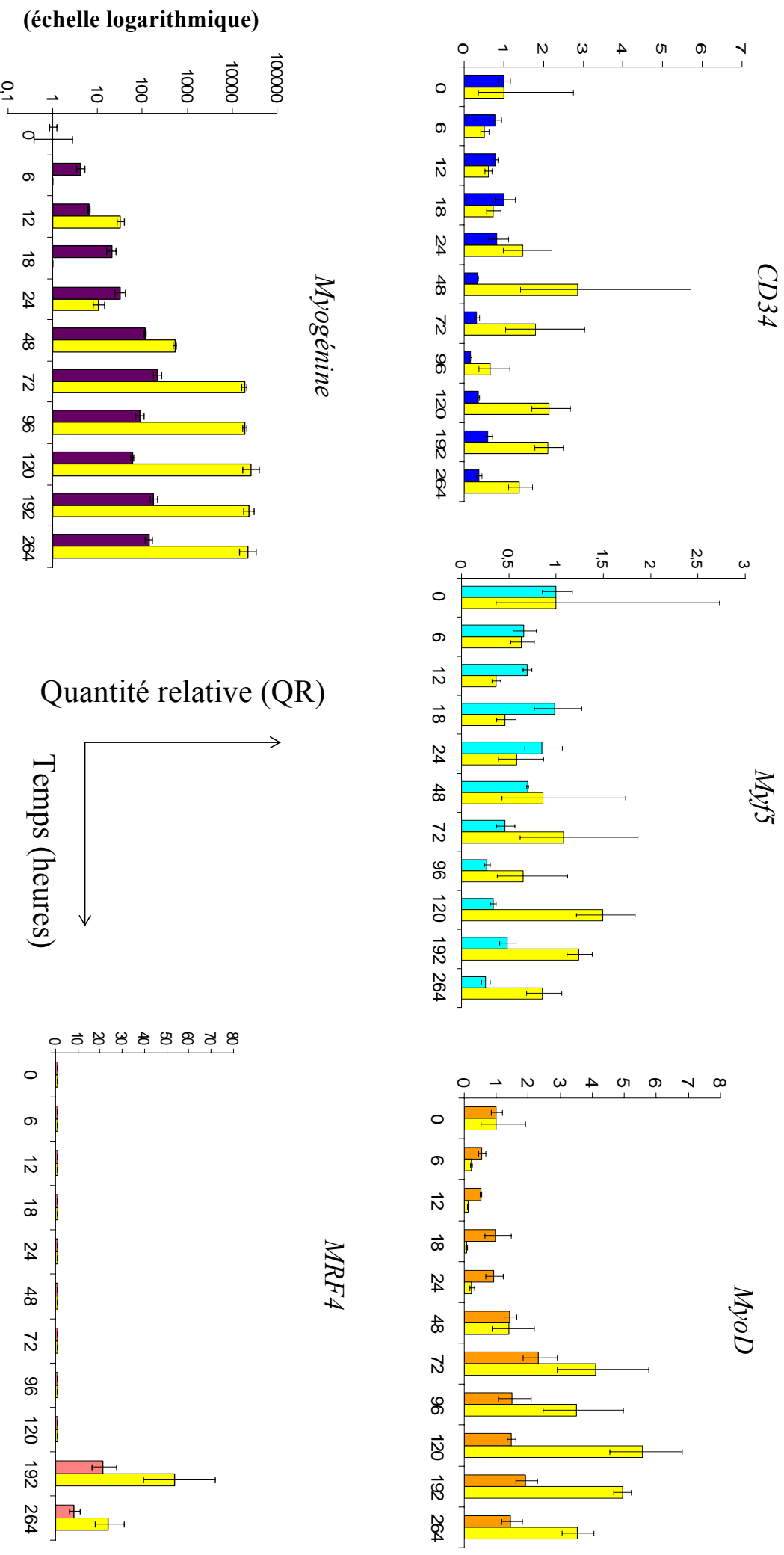


Figure 73 : Analyse en RT-PCR semi-quantitative du niveau d'expression du marqueur de prolifération CD34 et des quatre principaux facteurs bHLH intervenant au cours de la myogénèse, Myf5, MyoD, Myogénine et MRF4. Les graphiques présentent le niveau d'expression des transcrits (en quantité relative d'ARN présent) en fonction du temps (en heures) après mise en milieu de différenciation (T0 à T264 heures). Cette quantité est représentée en jaune quand les cellules ont été transfectées par la construction pcDNA/TOPO3.1-Pofit1 murin, l'autre couleur désignant le niveau d'expression dans une cinétique de différenciation de cellules non transfectées.

Comme nous l'avons vu dans la partie 3 de l'exposé bibliographique, les facteurs *Myf5* et *MyoD* jouent non seulement un rôle dans la détermination des cellules souches en cellules myoblastiques, mais leur expression est également nécessaire pour la différenciation des myoblastes en myotubes. Au cours de la cinétique de différenciation des cellules témoins (non transfectées), les transcrits des gènes *Myf5* et *MyoD* sont présents de manière permanente (T0 à T264) (Fig.73.). Aucune variation significative de la quantité de transcrits *Myf5* n'est observée jusqu'à T48, par la suite celle-ci diminue jusqu'à la fin de la cinétique. Cette variation d'expression est en corrélation avec l'état différencié des cellules C2C12 à ces stades. La quantité de transcrits *MyoD* est relativement faible au début de la cinétique puis augmente légèrement à partir de T48. Dans les cellules transfectées, des transcrits des gènes *Myf5* et *MyoD* sont, comme précédemment, présents tout au long de l'étude. Cependant, le profil d'expression de *Myf5* est inversé. En effet, si en début de cinétique le taux d'expression est relativement comparable à celui des cellules non transfectées, il ne diminue pas mais augmente à partir de T48. De même, le taux d'expression des transcrits *MyoD*, plus faible que dans les cellules non transfectées en début de cinétique, augmente en fin de cinétique à partir de T72 au lieu de rester stable. Ces résultats peuvent être mis en relation avec le niveau d'expression de *Pofut1*. En effet, il est envisageable qu'en début de cinétique, quand les cellules transfectées expriment fortement *Pofut1*, les récepteurs Notch soient hyper-activés et inhibent donc la transcription du facteur *MyoD* (cf. partie 3, exposé bibliographique). Les répercussions de l'activité des récepteurs Notch sur *Myf5* ne sont pas déterminés, il est cependant envisageable que leur hyper-activation inhibe également ce facteur myogénique. A des temps plus avancés de la cinétique, l'expression de *Pofut1* redevenant basale, les cellules reprennent le processus de différenciation myogénique et pour ce faire expriment des taux importants d'ARNm *MyoD* et *Myf5* qui décroissent de nouveau en fin de cinétique.

Des transcrits correspondant aux gènes *Myogénine* et *MRF4* (*Myf6*) sont normalement présents tardivement au cours de la différenciation de cellules myoblastiques (Buckingham, 1992). Au cours de la cinétique de différenciation des cellules C2C12 non transfectées, le gène *Myogénine* est exprimé de façon basale dès T0. Par la suite, son expression augmente très rapidement, dès 6h de différenciation (T6), pour atteindre un maximum à T72. La quantité de transcrits reste importante jusqu'à la fin de la cinétique, en relation avec un nombre de myotubes important. Dans les cellules transfectées, l'expression du gène *Myogénine* est retardée puisque la quantité de transcrits ne devient significative qu'à partir de

T24. Toutefois, des ARNm *Myogénine* sont détectés à T12. Cette présence peut être liée à un artefact de manipulation dû à la présence dans cette boîte de culture d'un faible taux de cellules transfectées. D'autres analyses seront nécessaires pour en avoir la confirmation. Ces résultats sont en accord avec ceux obtenus avec les facteurs MyoD et Myf5. La surexpression de *Pofut1* entraîne donc un retard d'environ 24 heures dans la différenciation des cellules myoblastiques. Quand la surexpression de *Pofut1* est arrêtée, l'expression de la Myogénine augmente comme pour tenter de rattraper le retard de différenciation. La quantité de transcrits *Myogénine* est, en effet, deux fois plus importante dans les cellules transfectées que dans les cellules « normales » en fin de cinétique (à partir de T72).

Les transcrits du gène *MRF4* sont absents en début de différenciation pour les deux cinétiques. *MRF4* est en effet un facteur intervenant de façon tardive puisqu'il est nécessaire à la différenciation terminale des cellules musculaires (Valdez *et al.*, 2000). Des transcrits *MRF4* ne sont détectés qu'à partir de T192. Il est toutefois à noter que l'expression de ce gène en fin de cinétique est environ deux fois plus importante dans les cellules où le gène *Pofut1* a été surexprimé.

En conclusion, on peut dire que la surexpression du gène *Pofut1* dans les cellules C2C12 semble en retarder la différenciation. En effet, cette surexpression doit avoir un effet hyper-activateur des récepteurs Notch. En conséquence, un plus grand nombre de cellules reste dans un état indifférencié dans les boîtes de culture contenant les cellules transfectées. Pour tous les facteurs bHLH étudiés ici (*Myf5*, *MyoD*, *Myogénine* et *MRF4*), on remarque une plus faible expression par rapport à celle observée dans les cellules « normales » en début de cinétique, celle-ci peut-être mise en relation avec une forte activité des récepteurs Notch, qui inhibent le processus de différenciation (Kitzmann, Communication au Club Muscle et développement III, 16-17 Décembre 2004, Paris). De façon générale, l'expression de ces quatre facteurs augmente de façon considérable (environ 2 fois plus de transcrits que dans les cellules « normales ») à partir de T48, temps où l'expression de *Pofut1* redevient basale dans les cellules transfectées. On peut donc supposer que l'arrêt de la surexpression de *Pofut1* entraîne un retour à la normale de l'activité des récepteurs Notch. L'inhibition de l'expression des facteurs bHLH étant levée, ceux-ci vont fortement s'exprimer afin que les cellules qui étaient dans un état de quiescence puissent rapidement fusionner avec les myotubes préexistants.

3.2.3.2 Inhibition de l'expression de *Pofut1*

Les conditions de culture et d'établissement de la cinétique de différenciation cellulaire sont les mêmes que décrites précédemment. Les transfections sont effectuées avec le vecteur de shRNA, pSilencer, en utilisant le système de nucléoporation (Amaxa). Ces manipulations sont actuellement en cours au laboratoire. Une étude préliminaire en RT-PCR semi-quantitative, 24 heures après transfection des cellules C2C12, montre une expression des transcrits *Pofut1* diminuée d'un facteur 3 par rapport à des cellules non transfectées.

L'expression de l'enzyme *Pofut1* endogène n'étant pas détectée par western blot, nous ne pourrions pas suivre par immunodétection, la diminution du niveau d'expression de la protéine. Les niveaux d'expression des transcrits codant les facteurs myogéniques à domaine bHLH seront également analysés.

3.3 Matériels et Méthodes

REACTION DE PCR	217
METHODES RELATIVES AUX CELLULES	217
Extraction des cellules satellites bovines	217
Tri des myoblastes au FACS (Fluorescence Activated Cell Sorter)	218
Marquage avec l'anticorps anti-laminine	218
Lipofection des cellules en culture primaire	219
Analyse par cytométrie de flux.....	219
Culture des cellules myoblastiques murines C2C12	219
Nucléoporation transitoire des cellules C2C12 et différenciation.....	220
DOSAGE DES ARN	220
RT-PCR SEMI-QUANTITATIVE EN TEMPS REEL	220

Réaction de PCR

La technique de PCR permet l'amplification d'un fragment d'ADN double brin selon un mode exponentiel du nombre de copies. L'ADN est synthétisé par la polymérase de la bactérie thermophile *Thermus aquaticus* (Taq polymérase) par extension à partir de deux amorces oligonucléotidiques.

Les PCR sont réalisées dans un volume final de 50 μ L contenant 30ng d'ADN en moyenne (ou une colonie bactérienne pour les PCR de criblage), 15pmol de chaque oligonucléotide, 0,5U de Taq polymérase, 2mM de MgCl₂, 0,2mM de chaque désoxynucléotide triphosphate (dATP, dCTP, dGTP, dTTP) et 1X de tampon. La Taq polymérase utilisée dans ces travaux est la UptiTherm Taq Polymérase (Interchim, Montluçon, France). La réaction classique est constituée de trois étapes : dénaturation thermique (96°C, 4min), amplification de l'ADN pendant 35 cycles (96°C, 30s; Tm -2°C, 45s (température inférieure de 2°C par rapport à la température d'hybridation des amorces (MWG, Ebersberg, Allemagne); 72°C, 1min par Kb d'ADN à amplifier) et terminaison de la synthèse d'ADN (72°C, 5min).

Méthodes relatives aux cellules

Extraction des cellules satellites bovines

A l'aide de ciseaux et de pinces stériles, un morceau de tissu est prélevé au niveau de la cuisse d'un embryon bovin (muscle *Semitendinosus*), puis émincé finement dans du milieu de culture cellulaire. Le mélange est ensuite centrifugé (1200t/min, 5min). Le culot est repris dans 10mL de solution de pronase (milieu HamF12, 0,15% pronase (p/v), 20mM HEPES, 10% sérum de veau fœtal (FCS)). La digestion se déroule à 37°C pendant 2 heures avec aspiration/refoulement toutes les 10min. Le mélange est centrifugé (1200t/min, 5min), puis le culot est repris dans 10mL de Trypsine 1X et laissé 20min à 37°C. La réaction est stoppée par ajout d'1mL de FCS. Les cellules sont filtrées sur des membranes à mailles de 100 μ m puis 40 μ m, et centrifugées (1400t/min, 7min). Elles sont enfin mises en culture dans du DMEM (Dubelco's modified Eagle's Medium) complété avec 10% de FCS, 2mM de L-glutamine, 50U/mL de pénicilline et 50 μ g/mL de streptomycine (milieu complet) et placées dans un incubateur à 37°C et 5% de CO₂.

Tri des myoblastes au FACS (Fluorescence Activated Cell Sorter)

48 heures après la mise en culture, les cellules sont lavées par du PBS 1X (13,7mM NaCl, 0,27mM KCl, 0,43mM Na₂HPO₄-7H₂O et 0,14mM KH₂PO₄), après élimination du milieu de culture par aspiration. Elles sont ensuite décollées de leur support par action de la trypsine (PBS 1X, EDTA 1mM, trypsine 0,05% (p/v)), pendant 3min à 37°C. L'action de la trypsine est stoppée par addition de milieu DMEM complet. Une série d'aspiration/refoulement permet de décoller les cellules encore adhérentes. La totalité des cellules est récupérée par centrifugation (1200t/min, 5min). Elles sont reprises dans 20mL de PBS 1X, comptées, aliquotées par $5 \cdot 10^5$ (autant de tubes que de cellules désirées) et de nouveau centrifugées (1200t/min, 10min). Les culots sont repris dans 90µL de PBS 1X additionné de 10µL d'anticorps anti-CD56 (couplé PE) (Becton-Dickinson, Le Pont de Claix, France) ou d'anticorps isotypique (couplé FITC, témoin négatif) (Becton-Dickinson). L'incubation est de 20min à 4°C. Les cellules sont rincées à plusieurs reprises dans du PBS 1X. Après une centrifugation finale, elles sont reprises dans 2mL de PBS 1X. La fluorescence est analysée au FACS et les cellules reconnues par l'anticorps anti-CD56 (les myoblastes) sont triées stérilement par le cytomètre, puis remises en culture en milieu complet.

Pour la différenciation des myoblastes bovins en myotubes, il faut que les cellules soient entre 80 et 90% de confluence. On remplace alors le milieu de culture par du milieu de différenciation (DMEM complété avec 2% de sérum de cheval, 2mM de L-glutamine, 50U/mL de pénicilline et 50µg/mL de streptomycine, 0,001% (p/v) d'insuline et 0,0005% (p/v) d'apo-transferrine); les premiers myotubes sont observables au bout de 72 heures.

Marquage avec l'anticorps anti-laminine

Les myoblastes bovins sont mis en culture à une densité de 10^6 cellules, dans une boîte de 100mm de diamètre dont le fond a été recouvert de lamelles de microscopie, puis de gélatine. Quand les cellules sont entre 80 et 90% de confluence, le milieu de culture est aspiré et remplacé par du milieu de différenciation. Quand elles ont fusionné pour former des myotubes, le milieu de différenciation est aspiré et les cellules sont fixées avec de l'éthanol froid 100%, pendant 4min à 4°C. Elles sont ensuite lavées 5min à deux reprises avec du PBS 1X, puis perméabilisées 10min à température ambiante avec une solution de TritonX100 (PBS 1X, 1% TritonX100 (v/v)), suivi de deux nouveaux lavages de 5min au PBS 1X. La saturation a lieu 30min à température ambiante (PBS 1X, 10% FCS (v/v)). Les cellules sont ensuite

incubées avec l'anticorps primaire anti-laminine (Sigma, Saint Louis, USA) dilué au 1/400^e (PBS 1X, 10% FCS (v/v)), suivi de deux lavages de 5min au PBS 1X. L'incubation dans l'anticorps secondaire anti-IgG de souris couplé FITC dilué au 1/500^e (PBS 1X, 10% FCS (v/v)), a lieu à température ambiante pendant 30min. Les cellules sont rincées 2 fois 10min avec du PBS 1X. Les lamelles recouvrant le fond de la boîte de culture sont retournées à l'aide d'une pince sur des lames puis observées au microscope à épifluorescence.

Lipofection des cellules en culture primaire

Les cellules sontensemencées sur des boîtes de 100mm de diamètre à une densité de $7,5 \cdot 10^5$ et ceci, 18 heures avant d'être transfectées. Juste avant la transfection, le milieu de culture est aspiré, et les cellules placées dans du DMEM minimum, après un lavage au PBS 1X pour éliminer tout le sérum. Après incubation 45min à température ambiante, le mélange réactionnel (3 μ g d'ADN plasmidique, 18 μ L de Fugene 6 (Roche, Mannheim, Germany), DMEM minimum qsp 300 μ L) est ajouté aux cellules. Huit heures après transfection, le milieu de culture (DMEM minimum + lipofectant) est remplacé par du milieu de culture complet.

Analyse par cytométrie de flux

Les analyses ont été réalisées avec un appareil FacsVantage (Becton-Dickinson) équipé d'un laser argon (488nm). La fluorescence verte de la IGFP (vecteur pVITRO2) et rouge de la PE (anti-CD56) permettent le tri des myoblastes bovins qui sont collectés respectivement avec des filtres bandes-passantes de 530nm et de 610nm. Un cytogramme de diffusion de la lumière à angle droit (SSC) et à petit angle (FSC) permet d'éliminer les débris et les agrégats cellulaires. Les paramètres SSC et FSC sont respectivement corrélés à la structure et à la taille des cellules. Les résultats se présentent sous la forme d'un cytogramme représentant la taille des cellules en fonction de la fluorescence. Afin de déterminer le pourcentage de transfection, 20 000 cellules transfectées et vivantes font l'objet d'une analyse préliminaire. Si le taux de transfection est correct, l'ensemble des cellules transfectées est alors trié.

Culture des cellules myoblastiques murines C2C12

La culture cellulaire est réalisée en condition standard dans un incubateur à 37°C et 5% de CO₂. Les cellules sont mises en culture dans du DMEM complet (identique à celui

utilisé pour les myoblastes bovins). 10^6 cellules sontensemencées par flasques de 75cm^2 , 48 heures avant transfection..

Nucléoporation transitoire des cellules C2C12 et différenciation

Après élimination du milieu de culture par aspiration, les cellules C2C12 sont lavées par du PBS 1X. Elles sont ensuite décollées de leur support par action de la trypsine (PBS 1X, EDTA 1mM, trypsine 0,05% (p/v)), pendant 3min à 37°C . L'action de la trypsine est stoppée par addition de milieu DMEM complet. Une série d'aspiration/refoulement permet de décoller les cellules encore adhérentes; la totalité des cellules est récupérée par centrifugation (1200t/min, 5min). Elles sont reprises dans 20mL de PBS 1X et comptées. Les cellules sont ensuite aliquotées par 10^6 (autant de tubes que de transfections à réaliser), puis à nouveau centrifugées. Les culots de cellules C2C12 sont resuspendus dans $100\mu\text{L}$ de solution nucleofector Kit V et $2\mu\text{g}$ de plasmide purifié sont ajoutés. Le tout est transféré dans une cuvette stérile de nucléoporation, ensuite placée dans l'appareil Amaxa. Le programme utilisé est le B-32. Les cellules nucléoporées sont reprises dans 1mL de milieu de culture complet à 37°C , puis remises en culture dans des flasques de 25cm^2 . Vingt-quatre heures après nucléoporation, le milieu de culture est remplacé après aspiration par du milieu de différenciation (identique à celui utilisé pour la différenciation des cellules de culture primaire bovines).

Dosage des ARN

La qualité et la quantité des ARN totaux extraits des cellules C2C12 sont contrôlées grâce à la méthode développée par Agilent technologies (Massy, France). $1\mu\text{L}$ des échantillons d'ARN extraits avec le RNeasy Mini Kit (QIAGEN, Les Ulis, France) sont déposés sur une puce Agilent RNA 6000 Nano LabChip et analysés avec un Agilent 2100 Bioanalyzer.

RT-PCR semi-quantitative en temps réel

Après dosage des ARN totaux correspondant à chaque point des cinétiques de différenciation des cellules C2C12, $10\mu\text{g}$ d'ARN sont rétro-transcrits grâce à l'utilisation du kit « High capacity cDNA Archive Kit » (Applied Biosystems, Norwalk, USA). Pour la PCR semi-quantitative, 2ng d'ADNc sont utilisés par point, additionnés de $10\mu\text{L}$ de tampon TaqMan, $1\mu\text{L}$ de la sonde complémentaire au transcrit dont on souhaite étudier l'expression et

8,8 μL d'eau; tous les points sont réalisés en triplicat. La PCR est exécutée par l'appareil Applied Biosystems 7900HT : Fast real time PCR System (plaque 96(384)~1h30/PCR). Lors de la PCR en temps réel, les étapes d'amplification sont les mêmes que lors de la PCR classique, c'est à dire, dénaturation, hybridation et élongation. L'émission de la fluorescence à chaque cycle, induite par la libération du « Reporter » par la polymérase, permet de détecter et visualiser les amplicons. L'analyse de la courbe d'amplification va permettre de déterminer la quantité de cibles présentes en début de réaction. La ligne de seuil est positionnée manuellement au-dessus du bruit de fond et au niveau de la phase exponentielle. L'intersection de la courbe d'amplification et de la ligne de seuil désigne un cycle ou C_t (Threshold Cycle) par projection sur l'axe des abscisses (Fig.74.). Le C_t est inversement proportionnel au nombre de cibles amplifiées initialement présentes dans le milieu réactionnel. Ainsi, deux échantillons ayant un nombre de copies différents auront 2 C_t différents.



Figure 74 : Représentation d'une amplification par PCR en temps réel, d'un fragment d'ADN sur une échelle logarithmique.

L'axe des abscisses correspond au nombre de cycles (C_t) et l'axe des ordonnées à l'augmentation de la fluorescence (Delta Rn), directement proportionnelle à la quantité de produit amplifié.

L'un des intérêts de la PCR en temps réel est le contrôle du rendement d'amplification. Théoriquement, pour un rendement de 100%, le nombre de copies de la séquence cible double à chaque cycle. Cependant, il est rare d'obtenir un tel rendement. Ainsi, une relation permet d'obtenir le nombre de copies (X_n) à un cycle donné (n) à partir d'un

nombre de copies initiales (X_0) en tenant compte du rendement (R). $X_n = X_0 \times (1+R)^n$, R étant compris entre 0 et 1. Deux types de quantification sont réalisables en PCR quantitative. La quantification absolue, exprimée en nombre de copies, consiste à mesurer la quantité de cible par rapport à des étalons dont le nombre de copies est connue. L'autre méthode (utilisée dans notre étude) est appelée quantification relative (ou PCR semi-quantitative) et consiste à mettre en relation le nombre de copies de la cible avec le nombre de copies d'un autre gène utilisé comme calibrateur. Dans notre étude, cette normalisation des résultats est effectuée grâce à l'utilisation pour chacun des échantillons, d'une sonde complémentaire à l'ARN 18S. Les résultats finaux, calculés par le logiciel SDS 2.2.2 (Applied Biosystems), sont donnés en quantité relative (QR) d'ARN cible présent dans l'échantillon.

Les numéros d'accès des sondes Applied Biosystems que nous avons utilisées sont : *Pofut1* : Mm00475567_m1; *MyoD* : Mm00440387_m1; *Myf5* : Mm00435125_m1; *MRF4* (*Myf6*) : Mm00435126_m1; *Myogénine* : Mm00446194_m1; *CD34* : Mm00519283_m1. Le rapporteur utilisé est la 6-carboxyfluorescéine (FAM).

Conclusion et perspectives

L'équipe « Glycogénome » de l'Unité de Génétique Moléculaire Animale (UMR 1061 INRA-Université de Limoges), dans laquelle j'ai effectué mon doctorat, s'intéresse notamment aux rôles potentiels joués par les glycoconjugués au cours de la prolifération et de la différenciation de cellules musculaires.

Plusieurs données bibliographiques récentes soulignent l'importance fonctionnelle des protéoglycanes et des glycoprotéines dans la myogenèse et la physiologie du muscle squelettique. Par exemple, les récepteurs Notch qui sont des glycoprotéines impliquées dans l'initiation et la modulation des voies de différenciation embryonnaire, participent à la différenciation, *in vitro*, des cellules myoblastiques murines de la lignée C2C12. De plus, de nombreuses études démontrent que les cascades de signalisations intracellulaires déclenchées par l'interaction des récepteurs Notch avec leurs ligands, sont dépendantes de leurs états de *O*-fucosylation (Okajima, 2003).

L'objectif de mon travail de thèse a consisté, à caractériser les *O*-fucosyltransférases bovines à l'échelle des gènes (*Pofut1* et *Pofut2*), de leurs transcrits, et des protéines qu'ils codent, à étudier l'évolution des *O*-fucosyltransférases au niveau du règne Animal, et à appréhender le rôle potentiellement joué par l'enzyme *O*-fucosyltransférase 1, au cours de la différenciation *in vitro* de cellules myoblastiques.

Ces travaux de thèse m'ont conduit, dans un premier temps, à l'établissement des structures des gènes *Pofut1* et *Pofut2* bovins, à la caractérisation de cinq variants transcriptionnels pour chacun des deux gènes et à la détermination de leurs expressions tissulaires. Suite à cette étude, les enzymes bovines *Pofut1* et *Pofut2* ont été caractérisées et leurs expressions spatio-temporelles définies.

La structure du gène *Pofut1* bovin, localisé sur le chromosome 13, comporte au moins 10 exons répartis sur au moins 39,4 Kb, alors que les gènes homologues humain et murin comportent respectivement 8 et 7 exons, répartis sur au moins 30,8 et 24,5 Kb. La structure du gène *Pofut2* bovin, dont la localisation chromosomique est encore indéterminée, comporte au moins 9 exons, répartis sur au moins 13,9 Kb. Le gène homologue murin qui comporte le même nombre d'exons s'étend sur 10,3 Kb, alors que le gène humain constitué de 10 exons est long de plus de 24 Kb. Si on ne tient compte que des exons codant les formes actives de Pofut1 et Pofut2, le découpage exon/intron est conservé chez toutes les espèces de Vertébrés étudiées à ce jour. La présence des gènes *Pofut1* et *Pofut2*, en une seule copie, dans les génomes d'un très grand nombre d'espèces (Vertébrés et Invertébrés) suggère un rôle fondamental de la *O*-fucosylation. Une étude phylogénétique, sur des séquences complètes de Pofut appartenant à des espèces recouvrant une diversité significative du monde Animal, révèle que les gènes ancestraux *Pofut1* et *Pofut2* étaient très probablement des gènes morcelés, déjà présents chez l'ancêtre des Bilatériens. Pour les deux gènes, les positions introniques les plus conservées sont situées en 5' du gène (position de l'intron 1 pour *Pofut1* et positions des introns 2 et 3 pour *Pofut2*). Ce résultat n'est pas surprenant. En effet, les régions régulatrices d'un gène ne sont pas seulement situées dans la région promotrice en 5'UTR, on trouve fréquemment des sites de fixation à de nombreux facteurs de régulation de l'expression génique dans les premiers introns d'un gène, spécialement dans l'intron 1 (Hu et Leung, 2006). Il est donc logique que de la position de tels introns soit conservée au cours de l'évolution.

Dans les tissus bovins adultes analysés, seul un variant transcriptionnel de *Pofut1* (*Pofut1a*) et quatre des variants de *Pofut2* (*Pofut2a*, *b*, *d* et *e*) sont présents de façon ubiquitaire. La répartition des variants de *Pofut1* est plus large dans les tissus embryonnaires, ce qui suggère que leur rôle potentiel est plus important à un stade du développement où les récepteurs de la famille Notch sont très actifs (Shi et Stanley, 2006). Parmi tous les variants transcriptionnels mis en évidence, seuls les variants *Pofut1a* et *Pofut2a* codent des protéines ayant des activités *O*-fucosyltransférases (l'activité de Pofut2B n'étant pas significative *in vitro*). Des analyses par western blot de l'expression des protéines Pofut, dans 13 tissus adultes et 8 tissus provenant d'embryons bovins, révèlent la présence des protéines codées par ces deux variants dans tous les tissus à l'exception des muscles squelettiques adultes. En effet, l'anticorps anti-Pofut1 n'y révèle qu'une protéine d'environ 32 KDa (et 36 KDa pour l'onglet) contre 45 KDa pour la forme active Pofut1A diglycosylée. Cette protéine de faible

masse moléculaire ne résulte pas de la traduction de l'un des variants transcriptionnels mis en évidence et serait sans doute le résultat d'un clivage protéolytique spécifique des tissus musculaires. Tout comme pour Pofut1, la forme active de l'enzyme Pofut2 est absente des muscles squelettiques adultes. Dans ces tissus, l'anticorps anti-Pofut2 détecte une forme de très haute masse moléculaire, environ 100 KDa, comparée à la protéine Pofut2A active triglycosylée (~50 KDa). Ce marquage, qui ne peut correspondre à la traduction d'aucun des variants transcriptionnels caractérisés, est sans doute le résultat d'associations protéiques résistantes aux conditions dénaturantes de l'électrophorèse. Les marquages anti-Pofut1 et anti-Pofut2 donnent des résultats similaires en utilisant des tissus d'autres espèces (souris et porc). Les spécificités des anticorps étant affirmées, il devient intéressant d'identifier les protéines révélées dans les muscles squelettiques adultes. Les expériences d'immunoprécipitation réalisées avec l'anticorps anti-Pofut1 sur des extraits protéiques de lysats cellulaires de muscles bovins n'ont donné, pour l'instant, aucun résultat, sans doute dû à la faible expression de ces protéines. Il devient dans ces conditions nécessaire de mettre au point une technique d'extraction protéique qui exclut les protéines contractiles, ce qui à terme favorise l'interaction des anticorps avec leurs cibles. De plus, jusqu'à présent, les expériences d'immunoprécipitation ont été effectuées avec des anticorps (anti-Pofut1) en suspension. La construction de colonnes de chromatographie d'affinité avec les anticorps polyclonaux anti-Pofut1 et anti-Pofut2 greffés doit être envisagée pour améliorer l'immunoprécipitation. Les protéines ainsi purifiées devront être séparées par électrophorèse sur un gel dénaturant à 1 ou 2 dimensions, puis identifiées grâce à une analyse en spectrométrie de masse.

Les variants transcriptionnels, *Pofut1b* à *Pofut1e* et *Pofut2b* à *Pofut2e*, ne sont pas homologues à ceux décrits chez d'autres espèces Animales. Ils ont été mis en évidence de façon indépendante à la suite de différentes expériences de RT-PCR, effectuées sur plusieurs tissus provenant de trois adultes et trois embryons bovins de race Charolaise. On peut donc penser qu'ils ne résultent pas d'artéfacts de manipulation. Ces transcrits, qui ne semblent pas être traduits *in vivo* pourraient avoir des rôles importants au sein des cellules. En effet, il est désormais connu que plusieurs types d'ARN (ARN, ARNt, ARNr) ne codent pas pour des protéines, ce sont les ARN non codants ou ARNnc (Eddy, 2001). Plusieurs formes d'ARNnc ont été récemment découvertes sans pour autant que leurs rôles ne soient déterminés. Ils pourraient, par exemple, participer à la régulation du niveau d'expression des protéines codées par les gènes à partir desquels ils sont transcrits (Larreau *et al.*, 2004).

L'enzyme Pofut1 bovine possède deux sites de *N*-glycosylation : le premier N⁶⁵ est conservé chez tous les Bilatériens (sauf chez les Nématodes *C. elegans*, *C. briggsae* et *S. ratti*, et un Insecte *H. coagulata*); le second N¹⁶³ n'est présent que chez les Mammifères et l'Insecte *B. mori*. Dans une première étape, nous nous sommes attachés à montrer que l'enzyme bovine était bien diglycosylée. Puis, par des expériences de mutagenèse dirigée, nous avons montré que si l'occupation du deuxième site de *N*-glycosylation ne semble pas indispensable au maintien de l'activité enzymatique, le premier *N*-glycane est absolument nécessaire au repliement correct de la protéine. En effet, sans le premier site de *N*-glycosylation, l'enzyme Pofut1 est mal conformée et s'agrège dans le RE. Etant donné la conservation de ce premier site de *N*-glycosylation chez pratiquement toutes les espèces étudiées jusqu'à présent et la mauvaise conformation de l'enzyme dépourvue de ce *N*-glycane, il n'est pas étonnant que celle-ci ait perdu toute activité enzymatique. D'autres expériences de mutagenèse dirigée sur des acides aminés conservés entre les espèces, comme par exemple, ceux inclus dans les motifs I, II et III sont envisagées au laboratoire et pourront compléter les données sur la relation structure/fonction de la *O*-fucosyltransférase 1.

Plus récemment, nous avons commencé un travail visant à appréhender le rôle potentiel de l'enzyme Pofut1 au cours de la différenciation *in vitro* de cellules myoblastiques. Pour ceci, deux voies de recherche ont été engagées simultanément. D'une part, nous avons mis en place au laboratoire une culture primaire de myoblastes bovins. Il a donc fallu mettre au point et maîtriser les différentes étapes techniques comme le prélèvement du tissu musculaire d'embryons, l'isolement, la culture, le tri et la différenciation des myoblastes. D'autre part, nous avons débuté une étude similaire sur les cellules de la lignée myoblastique murine C2C12. Le but d'un tel dispositif est de mettre en évidence les répercussions d'une surexpression ou d'une inhibition de l'expression du gène *Pofut1* sur la différenciation *in vitro* de myoblastes en myotubes. Ces études reposent sur la transfection des cellules avec un vecteur d'expression eucaryote permettant la surexpression transitoire de l'enzyme Pofut1 (bovine ou murine selon l'origine des myoblastes) ou avec un vecteur de shRNA permettant la dégradation des transcrits *Pofut1*. Elles nécessitent la mise en place de cinétiques de différenciation de cellules transfectées et non transfectées. Des expériences de RT-PCR semi-quantitative et de western blot permettent d'analyser les niveaux d'expression des transcrits et des protéines, pour *Pofut1* et pour des marqueurs myogéniques, en l'occurrence les principaux membres de la famille de facteurs de transcription à domaine bHLH (*Myf5*, *MyoD*, *Myogénine* et *MRF4*).

Etant donné les nombreuses difficultés rencontrées pour transfecter efficacement les cellules bovines en culture primaire, leur utilisation a été retardée. Ainsi, nous avons décidé d'opérer les modulations d'expression de *Pofut1* sur la lignée cellulaire murine. Ce volet de recherche se poursuit à l'heure actuelle au sein de notre équipe. La transfection des cellules C2C12 avec un vecteur de surexpression contenant l'ORF du gène murin *Pofut1*, semble entraîner un retard dans le processus de différenciation myogénique. En effet, durant les 72 premières heures après induction de la différenciation, les profils d'expression des facteurs bHLH dans les cellules surexprimant *Pofut1* sont retardés par rapport aux cellules témoins. Quand l'expression de *Pofut1* revient à un niveau basal, ces facteurs sont de nouveau exprimés et ce, de façon beaucoup plus importante que dans des cellules témoins. Ces résultats peuvent être rapprochés de ce qui est connu à l'heure actuelle du rôle des récepteurs Notch. En effet, nous savons que les interactions entre Notch et ses ligands ayant lieu dans la matrice extracellulaire sont modulées par leur état de *O*-fucosylation. Au cours de la myogenèse, le récepteur Notch a pour fonction principale de maintenir un pool de cellules dans un état indifférencié. Il est donc envisageable qu'une surexpression de *Pofut1* dans les myoblastes provoque une hyper-activation du récepteur Notch, empêchant ainsi les cellules transfectées de se différencier. Quand le niveau d'expression de *Pofut1* retrouve son niveau basal, les cellules sont de nouveau engagées dans une voie de différenciation myogénique et expriment de fortes quantités de facteurs bHLH, afin de fusionner avec les myotubes préexistants. Ceci est en accord avec les niveaux d'expression des transcrits *CD34*. En effet, les cellules transfectées expriment plus longtemps ce marqueur, en relation avec le maintien de l'état non fusionné des myoblastes.

Des expériences complémentaires sont évidemment nécessaires pour enrichir ces résultats. Ainsi, il est envisagé de vérifier par western blot le niveau d'expression de la protéine *Pofut1* au cours des cinétiques de différenciation. De même, il serait prudent de confirmer, à l'aide d'anticorps commerciaux dirigés contre les facteurs bHLH, les différences d'expression constatées au niveau transcriptionnel.

La *O*-fucosylation n'est pas la seule modification post-traductionnelle qui jouerait potentiellement un rôle dans la différenciation de cellules musculaires. Ainsi, deux autres approches dites « globales » sont actuellement développées au sein de l'équipe.

> La première consiste à étudier l'expression des principaux gènes de la glycosylation au cours de la différenciation des cellules musculaires de la lignée murine C2C12. Cette approche repose sur la mise en place et l'utilisation de la technologie TaqMAN Low Density micro-Array (TLDA), qui autorise l'analyse simultanée par PCR en temps réel, des niveaux d'expression de plusieurs centaines de gènes de la glycosylation. Cette analyse permettra de déterminer les principaux gènes qui influencent le déroulement du programme myogénique chez la souris. Sur les 375 gènes de la glycosylation analysés (ce qui représente une analyse transcriptomique de plus de 70% du "glycogénome"), les $\frac{3}{4}$ environ sont exprimés et 17% présentent des variations (d'un facteur supérieur ou égal à 4) de leurs quantités de transcrits au cours de la cinétique de différenciation. Quand les cellules C2C12 surexpriment *Pofut1*, 40 gènes de la glycosylation ont une expression différente par rapport à celle observée dans les cellules témoins. Les gènes exprimés différemment correspondraient à des gènes dont la régulation pourrait être dépendante d'un réseau moléculaire incluant potentiellement l'enzyme Pofut1.

> La deuxième, qui repose également sur la technologie TLDA consiste en l'étude des niveaux d'expression de gènes codant des protéines *O*-fucosylables (Notch et ses ligands, Cripto, etc.), des protéines impliquées dans les voies de signalisation de Notch (CBF1), les glycosyltransférases impliquées dans l'élongation du *O*-fucose lié (Fringes, β 4-galactosyltransférase 1) ou encore des protéines intervenant dans la régulation de la myogenèse (Wnts, BMP, TGF β). Cette étude, conduite en parallèle sur les cellules surexprimant *Pofut1* et les cellules témoins est actuellement en cours.

Jusqu'ici, c'est l'analyse des conséquences de la surexpression de *Pofut1* que nous avons entrepris. Des études similaires où le gène *Pofut1* est sous-exprimé débiteront prochainement. Des tests préliminaires avec le vecteur de shRNA dont nous disposons montrent que celui-ci est fonctionnel puisque 24 heures après transfection des cellules C2C12, l'expression des transcrits *Pofut1* est diminuée d'un facteur 3 par rapport à celle des cellules non transfectées.

Les résultats obtenus dans le cadre de ce travail de thèse apportent de nouvelles connaissances sur la *O*-fucosylation en général et sur les gènes de *O*-fucosyltransférases en particulier. Une étude préliminaire a par ailleurs ouvert une nouvelle voie de recherche visant à comprendre le rôle de l'enzyme Pofut1, au cours d'un processus développemental complexe, la Myogenèse.

Nouveau chapitre de la thèse
Mentor : Jacques Christen

1 Place et enjeux de la thèse

Il y a deux critères majeurs dans la sélection de races bovines destinées à la consommation, la qualité de leur viande et la quantité de muscle. Tous les consommateurs de viande savent en apprécier la qualité. Celle-ci est une propriété globale résultant de critères tels que tendreté, jutosité, teneur en gras intramusculaire, saveur et couleur du muscle. Pour des raisons de productivité, les éleveurs de bovins allaitants sont amenés à sélectionner génétiquement les animaux ayant une forte musculature, une caractéristique qui se décide dès les premières semaines de la vie, et qui fait appel à des mécanismes complexes gouvernant la différenciation et la croissance du tissu musculaire.

Au cours du développement embryonnaire des Vertébrés, certaines cellules dites somitiques vont se différencier en cellules musculaires, c'est-à-dire en myoblastes puis en myotubes. Cette différenciation se fait grâce à l'influence des cellules avoisinantes qui émettent ou portent à leur surface des protéines. Celles-ci vont spécifier aux cellules en différenciation, elles-mêmes porteuses de protéines dites réceptrices, l'évolution qu'elles doivent suivre. La myogenèse ou différenciation des cellules musculaires est un processus hautement régulé. Dans le cadre d'un travail sur les animaux d'élevage destinés à la consommation, la compréhension de ces mécanismes de régulation est essentielle, puisqu'ils conditionnent le nombre et la taille des myotubes, donc la masse musculaire des animaux. Les interactions protéiques au sein de la matrice extra-cellulaire sont souvent modulées par les modifications post-traductionnelles des protéines et, entre autres, par leur état de glycosylation. Dans les interactions protéiques déclenchant les voies de signalisation myogénique, le *O*-fucose joue un rôle majeur. La *O*-fucosylation est un processus qui consiste à greffer un fucose sur un résidu séryl ou thréonyl compris dans des motifs peptidiques consensus. Le fucose peut rester seul ou être allongé en oligosaccharide par d'autres enzymes de la glycosylation.

Dans le cadre de ma thèse, je m'intéresse aux gènes codant les enzymes responsables de la *O*-fucosylation chez le bovin, *pofut1* et *pofut2*. D'un point de vue fondamental, je m'efforce de comprendre comment fonctionnent ces gènes dans la cellule, et plus particulièrement dans la cellule musculaire. L'identification et la caractérisation de ces deux gènes chez le bovin, ainsi que la compréhension de leur rôle au cours de la différenciation de cellules musculaires en culture, permettraient de mettre en corrélation *O*-fucosylation et masse musculaire et donc à terme, une sélection génétique des bovins pour une meilleure maîtrise de la quantité de viande. Il faudra évidemment de nombreuses années pour arriver à de tels résultats; mon travail ne représente que les prémices du projet. Je peux néanmoins esquisser quelques conséquences indésirables d'une telle sélection génétique. En effet, une hypermusculature pourrait être accompagnée d'une moins bonne qualité de la viande et d'une quantité de gras insuffisante pour garder ses propriétés (conservation, cuisson, goût). Une autre étudiante du laboratoire s'efforce d'ailleurs de comprendre les mécanismes conduisant à la présence, plus ou moins importante, de gras dans le muscle chez différentes races bovines. En fonction des pays, les gens ne consomment pas le même type de viande. En effet, au Japon par exemple, on la consomme extrêmement grasse. La viande de bovins hypermusclés ne pourra donc pas être exportée dans n'importe quels pays; il faudra éventuellement effectuer une étude de marché avant d'envisager une commercialisation à grande échelle.

La *O*-fucosylation est depuis quelques années un sujet prisé internationalement. Par exemple, plusieurs équipes américaines travaillent sur la caractérisation des gènes responsables de cette fonction, chez l'Homme, la souris ou encore la drosophile. Ainsi, elles ont pu montrer que le gène *pofut1* était absolument nécessaire au développement d'un embryon de souris. En effet, si on supprime ce gène, l'embryon meurt au milieu de la gestation avec de nombreux problèmes développementaux touchant à la fois les tissus cardiaques, nerveux et musculaires. Néanmoins, ces équipes n'effectuent pas des études identiques aux miennes puisqu'elles s'appliquent surtout à comprendre de façon globale le rôle de la *O*-fucosylation et surtout son éventuelle intervention dans certaines pathologies. Au cours de congrès internationaux, j'ai pu discuter avec les directeurs de ces grands laboratoires et établir plusieurs collaborations qu'il m'a fallu gérer. J'ai pu échanger avec ces personnes du matériel biologique nécessaire à nos études respectives. Il faut, dans ce genre de cas, se tenir au courant mutuellement de l'avancement des travaux et savoir partager ses résultats expérimentaux.

J'ai réalisé ma thèse au sein de l'UMR 1061 INRA/Université de Limoges sur le site de la Faculté des Sciences et Techniques. Cette Unité de Génétique Moléculaire Animale est constituée de trois équipes : l'équipe « Génomique et traçabilité », l'équipe « Génomique et muscle » et l'équipe « Glycogénome ». Mon travail s'inscrit dans le cadre de cette dernière équipe. Plusieurs critères ont motivé mon choix d'effectuer ma thèse au sein de ce laboratoire. Le fait de travailler sur la *O*-fucosylation a suscité mon intérêt, puis après avoir fait une étude bibliographique importante, j'ai largement participé avec mes directeurs de thèse à l'élaboration du sujet. Le laboratoire, étant associé à l'INRA, offre, outre un support scientifique évident, des équipements nombreux permettant des apprentissages techniques divers. De plus, l'équipe glycogénome fait partie d'un groupement national de recherche sur les glycosyltransférases, le G3 rattaché au CNRS, qui se réunit deux fois par an, ce qui facilite les échanges entre chercheurs et par la même, les collaborations.

2 Gestion du projet et moyens mis en œuvre

A mon arrivée en thèse, la thématique dans laquelle s'intégrait mon sujet débutait. Après avoir réalisé une importante étude bibliographique, nous avons mis en place un sujet bien défini. Une allocation MRT, des vacances en enseignement ainsi qu'un emploi d'ATER (Attaché Temporaire d'Enseignement et de Recherche) au sein de la Faculté des Sciences et Techniques m'ont permis d'être financée durant quatre années.

Mon groupe de projet se constituait de mes directeurs de thèse M. Maftah Abderrahman, professeur de l'Université de Limoges et M^{elle} Germot Agnès, Maître de conférences de l'Université de Limoges. J'ai également bénéficié de l'aide précieuse de M. Dupuy Fabrice, maître de conférences de l'Université de Limoges. Ce groupe se réunissait régulièrement pour organiser, réorienter le projet mais également me conseiller à tous points de vue lorsque je le sollicitais. J'ai également pu bénéficier de l'aide technique plus ponctuelle d'une technicienne, de deux ingénieurs ainsi que de celle des autres membres du laboratoire.

Etant doctorante au sein d'une unité mixte de recherche INRA, j'ai bénéficié de l'appui d'un comité de thèse constitué de chercheurs d'horizons divers, chargé de suivre l'évolution de mes travaux et me conseiller lors de nos rencontres annuelles. Plusieurs fois par an, j'ai présenté mon travail à tous les membres du laboratoire ainsi que lors de différents congrès et séminaires ce qui m'a permis d'acquérir une aisance certaine en communication orale.

Le matériel utilisé a été principalement financé par l'INRA, l'Université de Limoges et le Conseil Régional du Limousin. J'ai pu disposer de l'ensemble des moyens techniques du laboratoire pour mener à bien mon projet. Cependant, aucun achat, s'il était justifié, ne m'a été refusé. Le tableau ci-dessous récapitule les dépenses directement associées à mon projet de thèse.

Catégorie		Dépenses
Ressources Humaines	Bourse et Salaires (et primes d'enseignements)	80 000 €
	Encadrants	40 000 €
	Permanents du laboratoire	10 000 €
	Administration	600 €
	Ecole doctorale	800 €
	Sous-total	131 400 €
Matériels	Consommables	45 000 €
	Equipements	4 000 €
	Sous-total	49 000 €
Infrastructures	Amortissement des locaux et aménagement	15 000 €
	Frais d'entretien logistique et maintenance	5 000 €
	Sous-total	20 000 €
Sous-traitance	Cytométrie en flux	3000 €
Documentation		1000 €
Déplacements		3000 €
Total		207 400 €

Le coût de ma thèse s'élève à près de 207 400 €.

3 Compétences développées dans le cadre du projet

Au cours de mon doctorat, j'ai développé de nombreuses compétences scientifiques et techniques. D'autre part, j'ai vu ma personnalité évoluer considérablement pendant ces quatre années. En effet, la thèse ne permet pas seulement l'acquisition de compétences strictement professionnelles mais aboutit très sûrement au développement individuel.

D'un point de vue scientifique, je me suis complètement appropriée le sujet, en acquérant une complète autonomie pour le mener à son terme. J'ai appris à organiser mon travail et à gérer le temps dont je disposais pour mener à bien mon projet. Mes encadrants m'ont toujours accordé leur confiance pour la gestion du matériel dont j'avais besoin. D'un point de vue technique, j'ai développé de larges compétences dans le domaine de la génétique et de la biologie moléculaire et cellulaire.

Les nombreuses rencontres avec mon groupe de projet m'ont permis d'apprendre à organiser et à conduire une réunion. Ces années m'ont amené à améliorer ma connaissance des outils informatiques en général, ma capacité rédactionnelle, mon esprit d'analyse et de synthèse mais aussi mon esprit critique. Ma participation à des congrès nationaux et internationaux a révélé, au cours de communications écrites ou orales, une réelle éloquence en français comme en anglais.

Tout au long de mon doctorat, le travail en équipe fut d'une importance capitale. J'ai largement participé à l'encadrement de plusieurs stagiaires de licence et de maîtrise, deux étudiantes en Master recherche et un doctorant arrivé dans l'équipe durant ma dernière année de thèse. L'enseignement que j'ai effectué à la Faculté des Sciences et Techniques de Limoges en tant que vacataire puis ATER a considérablement développé mes qualités pédagogiques et mon aptitude à être le leader d'un groupe de travail. Ce fut une expérience extrêmement enrichissante notamment lorsque j'eus l'occasion d'enseigner dans des domaines peu familiers.

La participation à de nombreuses formations dispensées grâce à l'école doctorale (sensibilisation à l'économie d'entreprise INECO, management de groupe, conduite de réunion, conduite de projet, premiers secours, semaine intensive d'anglais, TOIEC) m'a fait découvrir d'autres domaines et a enrichi mes connaissances et compétences.

4 Evaluation des retombées de la thèse

Durant la conduite de ce projet, j'ai développé diverses compétences techniques et scientifiques qui me permettront de poursuivre ma carrière dans le monde de la recherche. De plus, durant ma thèse, j'ai pu mettre au point des techniques d'isolement et de culture de cellules musculaires bovines. Celles-ci pourront servir à d'autres membres du laboratoire, étudiants ou permanents, pour étudier d'autres gènes impliqués dans les processus myogéniques. J'ai aussi dû créer des outils nécessaires à mon travail, comme par exemple des anticorps dirigés contre les protéines auxquelles je me suis intéressée. Ces anticorps pourront être utilisés par les étudiants qui prendront la suite de ce projet. Les collaborations que j'ai établies avec les directeurs de laboratoires aux USA continueront après mon départ et contribueront, je l'espère, à des progrès significatifs dans le domaine de la *O*-fucosylation.

Mon travail aura permis de faire un état des lieux des connaissances sur ce sujet. De plus, la *O*-fucosylation s'exerce sur des protéines dont le rôle est fondamental au cours du développement embryonnaire et dont le dysfonctionnement entraîne de nombreuses pathologies. Une meilleure compréhension de ces mécanismes chez le bovin, permettrait l'ouverture de nouvelles pistes de recherche chez l'Homme.

5 Conclusion

Je considère que les doctorants d'aujourd'hui plus que ceux d'hier doivent attacher une importance particulière à préciser tout au long de leur thèse leur projet professionnel et à privilégier leur ouverture d'esprit. C'est ce que je me suis attachée à faire tout au long de ces années. Pour moi, la thèse est à la fois une expérience professionnelle et une formation offrant un éventail assez large de possibilités. Personnellement, j'ai choisi de continuer dans le monde de la recherche, par la réalisation d'un stage post-doctoral. Ayant réalisé toutes mes études à Limoges, je pense qu'il est important que je finalise mon expérience dans un environnement différent afin de parfaire mes compétences et d'en découvrir de nouvelles. Tout ceci apportera également une valeur supplémentaire à mon projet professionnel. Par la suite, je souhaiterais intégrer le monde de la recherche publique ou privée. Dans tous les cas, mon objectif premier est d'évoluer dans un domaine professionnel passionnant avec tous les moyens humains, techniques et financiers nécessaires à la valorisation globale de tout individu.

Références bibliographiques

- Abeijon, C., and Hirschberg, C. B. (1992) Topography of glycosylation reactions in the endoplasmic reticulum. *Trends Biochem Sci* **17**, 32-6.
- Abeijon, C., Mandon, E. C., and Hirschberg, C. B. (1997) Transporters of nucleotide sugars, nucleotide sulfate and ATP in the Golgi apparatus. *Trends Biochem Sci* **22**, 203-7.
- Adams, J. C., and Tucker, R. P. (2000) The thrombospondin type 1 repeat (TSR) superfamily: diverse proteins with related roles in neuronal development. *Dev Dyn* **218**, 280-99.
- Akimoto, Y., Kreppel, L. K., Hirano, H., and Hart, G. W. (1999) Localization of the O-linked N-acetylglucosamine transferase in rat pancreas. *Diabetes* **48**, 2407-13.
- Anderson, D. C., and Springer, T. A. (1987) Leukocyte adhesion deficiency: an inherited defect in the Mac-1, LFA-1, and p150,95 glycoproteins. *Annu Rev Med* **38**, 175-94.
- Annaert, W., and De Strooper, B. (1999) Presenilins: molecular switches between proteolysis and signal transduction. *Trends Neurosci* **22**, 439-43.
- Arsenian, S., Weinhold, B., Oelgeschlager, M., Ruther, U., and Nordheim, A. (1998) Serum response factor is essential for mesoderm formation during mouse embryogenesis. *Embo J* **17**, 6289-99.
- Artavanis-Tsakonas, S. (1997) Alagille syndrome--a notch up for the Notch receptor. *Nat Genet* **16**, 212-3.
- Artavanis-Tsakonas, S., Rand, M. D., and Lake, R. J. (1999) Notch signaling: cell fate control and signal integration in development. *Science* **284**, 770-6.
- Asano, M., Furukawa, K., Kido, M., Matsumoto, S., Umesaki, Y., Kochibe, N., and Iwakura, Y. (1997) Growth retardation and early death of beta-1,4-galactosyltransferase knockout mice with augmented proliferation and abnormal differentiation of epithelial cells. *Embo J* **16**, 1850-7.
- Asch, A. S., Silbiger, S., Heimer, E., and Nachman, R. L. (1992) Thrombospondin sequence motif (CSVTCG) is responsible for CD36 binding. *Biochem Biophys Res Commun* **182**, 1208-17.
- Baboval, T., Koul, O., and Smith, F. I. (2000) N-glycosylation site occupancy of rat alpha-1,3-fucosyltransferase IV and the effect of glycosylation on enzymatic activity. *Biochim Biophys Acta* **1475**, 383-9.
- Bale, M. D. (1987) Noncovalent and covalent interactions of thrombospondin with polymerizing fibrin. *Semin Thromb Hemost* **13**, 326-34.
- Baron, M. (2003) An overview of the Notch signalling pathway. *Semin Cell Dev Biol* **14**, 113-9.
- Barrantes, I. B., Elia, A. J., Wunsch, K., Hrabe de Angelis, M. H., Mak, T. W., Rossant, J.,

- Conlon, R. A., Gossler, A., and de la Pompa, J. L. (1999) Interaction between Notch signalling and Lunatic fringe during somite boundary formation in the mouse. *Curr Biol* **9**, 470-80.
- Barreaud, J. P., Saunier, K., Souchaire, J., Delourme, D., Oulmouden, A., Oriol, R., Leveziel, H., Julien, R., and Petit, J. M. (2000) Three bovine alpha2-fucosyltransferase genes encode enzymes that preferentially transfer fucose on Galbeta1-3GalNAc acceptor substrates. *Glycobiology* **10**, 611-21.
- Barret, A., Forestier, L., Deslys, J. P., Julien, R., and Gallet, P. F. (2005) Glycosylation-related gene expression in prion diseases: PrPSc accumulation in scrapie infected GT1 cells depends on beta-1,4-linked GalNAc-4-SO4 hyposulfation. *J Biol Chem* **280**, 10516-23.
- Bartoli, M., and Richard, I. (2005) Calpains in muscle wasting. *Int J Biochem Cell Biol* **37**, 2115-33.
- Beauchamp, J. R., Heslop, L., Yu, D. S., Tajbakhsh, S., Kelly, R. G., Wernig, A., Buckingham, M. E., Partridge, T. A., and Zammit, P. S. (2000) Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *J Cell Biol* **151**, 1221-34.
- Becker, D. J., and Lowe, J. B. (2003) Fucose: biosynthesis and biological function in mammals. *Glycobiology* **13**, 41R-53R.
- Bjoern, S., Foster, D. C., Thim, L., Wiberg, F. C., Christensen, M., Komiyama, Y., Pedersen, A. H., and Kisiel, W. (1991) Human plasma and recombinant factor VII. Characterization of O-glycosylations at serine residues 52 and 60 and effects of site-directed mutagenesis of serine 52 to alanine. *J Biol Chem* **266**, 11051-7.
- Black, B. L., Martin, J. F., and Olson, E. N. (1995) The mouse MRF4 promoter is trans-activated directly and indirectly by muscle-specific transcription factors. *J Biol Chem* **270**, 2889-92.
- Blackwell, T. K., and Weintraub, H. (1990) Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**, 1104-10.
- Blaumueller, C. M., and Artavanis-Tsakonas, S. (1997) Comparative aspects of Notch signaling in lower and higher eukaryotes. *Perspect Dev Neurobiol* **4**, 325-43.
- Bober, E., Franz, T., Arnold, H. H., Gruss, P., and Tremblay, P. (1994) Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development* **120**, 603-12.

- Brand-Saberi, B., and Christ, B. (1999) Genetic and epigenetic control of muscle development in vertebrates. *Cell Tissue Res* **296**, 199-212.
- Braun, T., Bober, E., and Arnold, H. H. (1992) Inhibition of muscle differentiation by the adenovirus E1a protein: repression of the transcriptional activating function of the HLH protein Myf-5. *Genes Dev* **6**, 888-902.
- Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E., and Arnold, H. H. (1989) A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. *Embo J* **8**, 701-9.
- Breton, C., and Imberty, A. (1999) Structure/function studies of glycosyltransferases. *Curr Opin Struct Biol* **9**, 563-71.
- Breton, C., Oriol, R., and Imberty, A. (1998) Conserved structural features in eukaryotic and prokaryotic fucosyltransferases. *Glycobiology* **8**, 87-94.
- Brockhausen, I., Yang, J. M., Burchell, J., Whitehouse, C., and Taylor-Papadimitriou, J. (1995) Mechanisms underlying aberrant glycosylation of MUC1 mucin in breast cancer cells. *Eur J Biochem* **233**, 607-17.
- Bruckner, K., Perez, L., Clausen, H., and Cohen, S. (2000) Glycosyltransferase activity of Fringe modulates Notch-Delta interactions. *Nature* **406**, 411-5.
- Buchberger, A., Ragge, K., and Arnold, H. H. (1994) The myogenin gene is activated during myocyte differentiation by pre-existing, not newly synthesized transcription factor MEF-2. *J Biol Chem* **269**, 17289-96.
- Buckingham, M. (1992) Making muscle in mammals. *Trends Genet* **8**, 144-8.
- Buckingham, M. (2001) Skeletal muscle formation in vertebrates. *Curr Opin Genet Dev* **11**, 440-8.
- Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meilhac, S., Montarras, D., Rocancourt, D., and Relaix, F. (2003) The formation of skeletal muscle: from somite to limb. *J Anat* **202**, 59-68.
- Byrd, J. C., and Bresalier, R. S. (2004) Mucins and mucin binding proteins in colorectal cancer. *Cancer Metastasis Rev* **23**, 77-99.
- Cacan, R., and Verbert, A. (1999) Free and N-linked oligomannosides as markers of the quality control of newly synthesized glycoproteins. *Biochem Biophys Res Commun* **258**, 1-5.
- Carnac, G., Primig, M., Kitzmann, M., Chafey, P., Tuil, D., Lamb, N., and Fernandez, A. (1998) RhoA GTPase and serum response factor control selectively the expression of MyoD without affecting Myf5 in mouse myoblasts. *Mol Biol Cell* **9**, 1891-902.

- Carraway, K. L., and Hull, S. R. (1991) Cell surface mucin-type glycoproteins and mucin-like domains. *Glycobiology* **1**, 131-8.
- Cates, G. A., Brickenden, A. M., and Sanwal, B. D. (1984) Possible involvement of a cell surface glycoprotein in the differentiation of skeletal myoblasts. *J Biol Chem* **259**, 2646-50.
- Chai, J., and Tarnawski, A. S. (2002) Serum response factor: discovery, biochemistry, biological roles and implications for tissue injury healing. *J Physiol Pharmacol* **53**, 147-57.
- Chai, W., Yuen, C. T., Kogelberg, H., Carruthers, R. A., Margolis, R. U., Feizi, T., and Lawson, A. M. (1999) High prevalence of 2-mono- and 2,6-di-substituted manol-terminating sequences among O-glycans released from brain glycopeptides by reductive alkaline hydrolysis. *Eur J Biochem* **263**, 879-88.
- Chapman, B. S., Eckart, M. R., Kaufman, S. E., and Lapointe, G. R. (1996) O-linked oligosaccharide on the 75-kDa neurotrophin receptor. *J Neurochem* **66**, 1707-16.
- Chazalet, V., Uehara, K., Geremia, R. A., and Breton, C. (2001) Identification of essential amino acids in the Azorhizobium caulinodans fucosyltransferase NodZ. *J Bacteriol* **183**, 7067-75.
- Chen, J., Lu, L., Shi, S., and Stanley, P. (2006) Expression of Notch signaling pathway genes in mouse embryos lacking beta4galactosyltransferase-1. *Gene Expr Patterns* **6**, 376-82.
- Chen, J., Moloney, D. J., and Stanley, P. (2001) Fringe modulation of Jagged1-induced Notch signaling requires the action of beta 4galactosyltransferase-1. *Proc Natl Acad Sci U S A* **98**, 13716-21.
- Chen, Y. J., Wing, D. R., Guile, G. R., Dwek, R. A., Harvey, D. J., and Zamze, S. (1998) Neutral N-glycans in adult rat brain tissue--complete characterisation reveals fucosylated hybrid and complex structures. *Eur J Biochem* **251**, 691-703.
- Chiba, A., Matsumura, K., Yamada, H., Inazu, T., Shimizu, T., Kusunoki, S., Kanazawa, I., Kobata, A., and Endo, T. (1997) Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve alpha-dystroglycan. The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin. *J Biol Chem* **272**, 2156-62.
- Christ, B., and Ordahl, C. P. (1995) Early stages of chick somite development. *Anat Embryol (Berl)* **191**, 381-96.
- Cieniewski-Bernard, C., Bastide, B., Lefebvre, T., Lemoine, J., Mounier, Y., and Michalski, J.

- C. (2004) Identification of O-linked N-acetylglucosamine proteins in rat skeletal muscle using two-dimensional gel electrophoresis and mass spectrometry. *Mol Cell Proteomics* **3**, 577-85.
- Cieniewski-Bernard, C., Mounier, Y., Michalski, J. C., and Bastide, B. (2005) O-GlcNAc level variations are associated with the development of skeletal muscle atrophy. *J Appl Physiol*.
- Cohen, B., Bashirullah, A., Dagnino, L., Campbell, C., Fisher, W. W., Leow, C. C., Whiting, E., Ryan, D., Zinyk, D., Boulianne, G., Hui, C. C., Gallie, B., Phillips, R. A., Lipshitz, H. D., and Egan, S. E. (1997) Fringe boundaries coincide with Notch-dependent patterning centres in mammals and alter Notch-dependent development in *Drosophila*. *Nat Genet* **16**, 283-8.
- Conboy, I. M., Conboy, M. J., Smythe, G. M., and Rando, T. A. (2003) Notch-mediated restoration of regenerative potential to aged muscle. *Science* **302**, 1575-7.
- Conlon, R. A., Reaume, A. G., and Rossant, J. (1995) Notch1 is required for the coordinate segmentation of somites. *Development* **121**, 1533-45.
- Cooper, H. S., Malecha, M. J., Bass, C., Fagel, P. L., and Steplewski, Z. (1991) Expression of blood group antigens H-2, Le(y), and sialylated-Le(a) in human colorectal carcinoma. An immunohistochemical study using double-labeling techniques. *Am J Pathol* **138**, 103-10.
- Cossu, G., and Borello, U. (1999) Wnt signaling and the activation of myogenesis in mammals. *Embo J* **18**, 6867-72.
- Cossu, G., Tajbakhsh, S., and Buckingham, M. (1996) How is myogenesis initiated in the embryo? *Trends Genet* **12**, 218-23.
- Costa, F. F. (2005) Non-coding RNAs: new players in eukaryotic biology. *Gene* **357**, 83-94.
- Cserjesi, P., Lilly, B., Hinkley, C., Perry, M., and Olson, E. N. (1994) Homeodomain protein MHox and MADS protein myocyte enhancer-binding factor-2 converge on a common element in the muscle creatine kinase enhancer. *J Biol Chem* **269**, 16740-5.
- Cui, X. Y., Hu, Q. D., Tekaya, M., Shimoda, Y., Ang, B. T., Nie, D. Y., Sun, L., Hu, W. P., Karsak, M., Duka, T., Takeda, Y., Ou, L. Y., Dawe, G. S., Yu, F. G., Ahmed, S., Jin, L. H., Schachner, M., Watanabe, K., Arsenijevic, Y., and Xiao, Z. C. (2004) NB-3/Notch1 pathway via Deltex1 promotes neural progenitor cell differentiation into oligodendrocytes. *J Biol Chem* **279**, 25858-65.
- Dahms, N. M., and Kornfeld, S. (1989) The cation-dependent mannose 6-phosphate receptor. Structural requirements for mannose 6-phosphate binding and oligomerization. *J Biol*

- Chem* **264**, 11458-67.
- Davis, J. S., Weakland, L. L., Farese, R. V., and West, L. A. (1987) Luteinizing hormone increases inositol trisphosphate and cytosolic free Ca²⁺ in isolated bovine luteal cells. *J Biol Chem* **262**, 8515-21.
- de Bernabe, D. B., van Bokhoven, H., van Beusekom, E., Van den Akker, W., Kant, S., Dobyens, W. B., Cormand, B., Currier, S., Hamel, B., Talim, B., Topaloglu, H., and Brunner, H. G. (2003) A homozygous nonsense mutation in the fukutin gene causes a Walker-Warburg syndrome phenotype. *J Med Genet* **40**, 845-8.
- Dennis, J. W., Granovsky, M., and Warren, C. E. (1999) Glycoprotein glycosylation and cancer progression. *Biochim Biophys Acta* **1473**, 21-34.
- Ding, J., Yang, L., Yan, Y. T., Chen, A., Desai, N., Wynshaw-Boris, A., and Shen, M. M. (1998) Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo. *Nature* **395**, 702-7.
- Dodou, E., Xu, S. M., and Black, B. L. (2003) mef2c is activated directly by myogenic basic helix-loop-helix proteins during skeletal muscle development in vivo. *Mech Dev* **120**, 1021-32.
- Doherty, D., Feger, G., Younger-Shepherd, S., Jan, L. Y., and Jan, Y. N. (1996) Delta is a ventral to dorsal signal complementary to Serrate, another Notch ligand, in Drosophila wing formation. *Genes Dev* **10**, 421-34.
- Dohrman, A., Miyata, S., Gallup, M., Li, J. D., Chapelin, C., Coste, A., Escudier, E., Nadel, J., and Basbaum, C. (1998) Mucin gene (MUC 2 and MUC 5AC) upregulation by Gram-positive and Gram-negative bacteria. *Biochim Biophys Acta* **1406**, 251-9.
- Dong, D. L., and Hart, G. W. (1994) Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol. *J Biol Chem* **269**, 19321-30.
- Dupuy, F., Germot, A., Julien, R., and Maftah, A. (2004) Structure/function study of Lewis alpha3- and alpha3/4-fucosyltransferases: the alpha1,4 fucosylation requires an aromatic residue in the acceptor-binding domain. *Glycobiology* **14**, 347-56.
- Dupuy, F., Germot, A., Marena, M., Oriol, R., Blancher, A., Julien, R., and Maftah, A. (2002) Alpha1,4-fucosyltransferase activity: a significant function in the primate lineage has appeared twice independently. *Mol Biol Evol* **19**, 815-24.
- Dupuy, F., Petit, J. M., Mollicone, R., Oriol, R., Julien, R., and Maftah, A. (1999) A single amino acid in the hypervariable stem domain of vertebrate alpha1,3/1,4-fucosyltransferases determines the type 1/type 2 transfer. Characterization of acceptor

- substrate specificity of the lewis enzyme by site-directed mutagenesis. *J Biol Chem* **274**, 12257-62.
- Eckhardt, A. E., and Goldstein, I. J. (1983) Isolation and characterization of a family of alpha-D-galactosyl-containing glycopeptides from Ehrlich ascites tumor cells. *Biochemistry* **22**, 5290-7.
- Eddy, S. R. (2001) Non-coding RNA genes and the modern RNA world. *Nat Rev Genet* **2**, 919-29.
- Edmondson, D. G., Cheng, T. C., Cserjesi, P., Chakraborty, T., and Olson, E. N. (1992) Analysis of the myogenin promoter reveals an indirect pathway for positive autoregulation mediated by the muscle-specific enhancer factor MEF-2. *Mol Cell Biol* **12**, 3665-77.
- Edmondson, D. G., and Olson, E. N. (1990) A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. *Genes Dev* **4**, 1450.
- Eloy-Trinquet, S., and Nicolas, J. F. (2002) Cell coherence during production of the presomitic mesoderm and somitogenesis in the mouse embryo. *Development* **129**, 3609-19.
- Etzioni, A., Frydman, M., Pollack, S., Avidor, I., Phillips, M. L., Paulson, J. C., and Gershoni-Baruch, R. (1992) Brief report: recurrent severe infections caused by a novel leukocyte adhesion deficiency. *N Engl J Med* **327**, 1789-92.
- Etzioni, A., Sturla, L., Antonellis, A., Green, E. D., Gershoni-Baruch, R., Berninsone, P. M., Hirschberg, C. B., and Tonetti, M. (2002) Leukocyte adhesion deficiency (LAD) type II/carbohydrate deficient glycoprotein (CDG) IIc founder effect and genotype/phenotype correlation. *Am J Med Genet* **110**, 131-5.
- Fernandez, M. S., Dennis, J. E., Drushel, R. F., Carrino, D. A., Kimata, K., Yamagata, M., and Caplan, A. I. (1991) The dynamics of compartmentalization of embryonic muscle by extracellular matrix molecules. *Dev Biol* **147**, 46-61.
- Finne, J., Krusius, T., Margolis, R. K., and Margolis, R. U. (1979) Novel mannitol-containing oligosaccharides obtained by mild alkaline borohydride treatment of a chondroitin sulfate proteoglycan from brain. *J Biol Chem* **254**, 10295-300.
- Fleming, R. J. (1998) Structural conservation of Notch receptors and ligands. *Semin Cell Dev Biol* **9**, 599-607.
- Florini, J. R., Ewton, D. Z., and Magri, K. A. (1991) Hormones, growth factors, and myogenic differentiation. *Annu Rev Physiol* **53**, 201-16.

- Fondeur-Gelinotte, M., Lattard, V., Oriol, R., Mollicone, R., Jacquinet, J.C., Mulliert, G., Gulberti, S., Netter, P., Magdalou, J., Ouzzine, M., and Fournel-Gigleux, S. (2006) Phylogenetic and mutational analyses reveal key residues for UDP-glucuronic acid binding and activity of beta1,3-glucuronosyltransferase I (GlcAT-I). *Protein Sci* **15**, 1667-1678
- Frazier, W. A. (1991) Thrombospondins. *Curr Opin Cell Biol* **3**, 792-9.
- Freeze, H. H., and Aebi, M. (2005) Altered glycan structures: the molecular basis of congenital disorders of glycosylation. *Curr Opin Struct Biol* **15**, 490-8.
- Gallagher, A., Aduse-Opoku, J., Rangarajan, M., Slaney, J. M., and Curtis, M. A. (2003) Glycosylation of the Arg-gingipains of *Porphyromonas gingivalis* and comparison with glycoconjugate structure and synthesis in other bacteria. *Curr Protein Pept Sci* **4**, 427-41.
- Geisow, M. J. (1991) Peptide synthesis--theme and variations. *Trends Biotechnol* **9**, 294-5.
- Gilfix, B. M., and Sanwal, B. D. (1980) Inhibition of myoblast fusion by tunicamycin and pantomycin. *Biochem Biophys Res Commun* **96**, 1184-91.
- Gohlke, M., Baude, G., Nuck, R., Grunow, D., Kannicht, C., Bringmann, P., Donner, P., and Reutter, W. (1996) O-linked L-fucose is present in *Desmodus rotundus* salivary plasminogen activator. *J Biol Chem* **271**, 7381-6.
- Gonzalez de Peredo, A., Klein, D., Macek, B., Hess, D., Peter-Katalinic, J., and Hofsteenge, J. (2002) C-mannosylation and o-fucosylation of thrombospondin type 1 repeats. *Mol Cell Proteomics* **1**, 11-8.
- Gossett, L. A., Kelvin, D. J., Sternberg, E. A., and Olson, E. N. (1989) A new myocyte-specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. *Mol Cell Biol* **9**, 5022-33.
- Gotting, C., Muller, S., Schottler, M., Schon, S., Prante, C., Brinkmann, T., Kuhn, J., and Kleesiek, K. (2004) Analysis of the DXD motifs in human xylosyltransferase I required for enzyme activity. *J Biol Chem* **279**, 42566-73.
- Greenwald, I. (1998) LIN-12/Notch signaling: lessons from worms and flies. *Genes Dev* **12**, 1751-62.
- Grewal, P. K., and Hewitt, J. E. (2003) Glycosylation defects: a new mechanism for muscular dystrophy? *Hum Mol Genet* **12 Spec No 2**, R259-64.
- Grewal, P. K., Holzfeind, P. J., Bittner, R. E., and Hewitt, J. E. (2001) Mutant glycosyltransferase and altered glycosylation of alpha-dystroglycan in the myodystrophy mouse. *Nat Genet* **28**, 151-4.

- Guerardel, Y., Chang, L. Y., Maes, E., Huang, C. J., and Khoo, K. H. (2006) Glycomic survey mapping of zebrafish identifies unique sialylation pattern. *Glycobiology* **16**, 244-57.
- Guernec, A., Berri, C., Chevalier, B., Wacrenier-Cere, N., Le Bihan-Duval, E., and Duclos, M. J. (2003) Muscle development, insulin-like growth factor-I and myostatin mRNA levels in chickens selected for increased breast muscle yield. *Growth Horm IGF Res* **13**, 8-18.
- Guidos, C. J. (2002) Notch signaling in lymphocyte development. *Semin Immunol* **14**, 395-404.
- Gupta-Bansal, R., Frederickson, R. C., and Brunden, K. R. (1995) Proteoglycan-mediated inhibition of A beta proteolysis. A potential cause of senile plaque accumulation. *J Biol Chem* **270**, 18666-71.
- Hallgren, P., Lundblad, A., and Svensson, S. (1975) A new type of carbohydrate-protein linkage in a glycopeptide from normal human urine. *J Biol Chem* **250**, 5312-4.
- Haltiwanger, R. S. (2002) Regulation of signal transduction pathways in development by glycosylation. *Curr Opin Struct Biol* **12**, 593-8.
- Haltiwanger, R. S., and Lowe, J. B. (2004) Role of glycosylation in development. *Annu Rev Biochem* **73**, 491-537.
- Hampton, R. Y. (2002) ER-associated degradation in protein quality control and cellular regulation. *Curr Opin Cell Biol* **14**, 476-82.
- Han, I., and Kudlow, J. E. (1997) Reduced O glycosylation of Sp1 is associated with increased proteasome susceptibility. *Mol Cell Biol* **17**, 2550-8.
- Hanover, J. A., Lai, Z., Lee, G., Lubas, W. A., and Sato, S. M. (1999) Elevated O-linked N-acetylglucosamine metabolism in pancreatic beta-cells. *Arch Biochem Biophys* **362**, 38-45.
- Harris, R. J., Leonard, C. K., Guzzetta, A. W., and Spellman, M. W. (1991) Tissue plasminogen activator has an O-linked fucose attached to threonine-61 in the epidermal growth factor domain. *Biochemistry* **30**, 2311-4.
- Harris, R. J., Ling, V. T., and Spellman, M. W. (1992) O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C. *J Biol Chem* **267**, 5102-7.
- Harris, R. J., and Spellman, M. W. (1993) O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology* **3**, 219-24.
- Harris, R. J., van Halbeek, H., Glushka, J., Basa, L. J., Ling, V. T., Smith, K. J., and Spellman, M. W. (1993) Identification and structural analysis of the tetrasaccharide

- NeuAc alpha(2-->6)Gal beta(1-->4)GlcNAc beta(1-->3)Fuc alpha 1-->O-linked to serine 61 of human factor IX. *Biochemistry* **32**, 6539-47.
- Hart, G. W. (1997) Dynamic O-linked glycosylation of nuclear and cytoskeletal proteins. *Annu Rev Biochem* **66**, 315-35.
- Hase, S., Kawabata, S., Nishimura, H., Takeya, H., Sueyoshi, T., Miyata, T., Iwanaga, S., Takao, T., Shimonishi, Y., and Ikenaka, T. (1988) A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX. *J Biochem (Tokyo)* **104**, 867-8.
- Hasegawa, H., Watanabe, M., Arisawa, Y., Teramoto, T., Kodaira, S., and Kitajima, M. (1993) Carbohydrate antigens and liver metastasis in colorectal cancer. *Jpn J Clin Oncol* **23**, 336-41.
- Hebert, D. N., Foellmer, B., and Helenius, A. (1995) Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. *Cell* **81**, 425-33.
- Helenius, A., and Aebi, M. (2001) Intracellular functions of N-linked glycans. *Science* **291**, 2364-9.
- Hennet, T., Hagen, F. K., Tabak, L. A., and Marth, J. D. (1995) T-cell-specific deletion of a polypeptide N-acetylgalactosaminyl-transferase gene by site-directed recombination. *Proc Natl Acad Sci U S A* **92**, 12070-4.
- Herscovics, A., and Orlean, P. (1993) Glycoprotein biosynthesis in yeast. *Faseb J* **7**, 540-50.
- Hickman, S., and Neufeld, E. F. (1972) A hypothesis for I-cell disease: defective hydrolases that do not enter lysosomes. *Biochem Biophys Res Commun* **49**, 992-9.
- Hofsteenge, J., Huwiler, K. G., Macek, B., Hess, D., Lawler, J., Mosher, D. F., and Peter-Katalinic, J. (2001) C-mannosylation and O-fucosylation of the thrombospondin type 1 module. *J Biol Chem* **276**, 6485-98.
- Holland, P. C., and Herscovics, A. (1986) Inhibition of myoblast fusion by the glucosidase inhibitor N-methyl-1-deoxynojirimycin, but not by the mannosidase inhibitor 1-deoxymannojirimycin. *Biochem J* **238**, 335-40.
- Holmes, E. H., Yen, T. Y., Thomas, S., Joshi, R., Nguyen, A., Long, T., Gallet, F., Maftah, A., Julien, R., and Macher, B. A. (2000) Human alpha 1,3/4 fucosyltransferases. Characterization of highly conserved cysteine residues and N-linked glycosylation sites. *J Biol Chem* **275**, 24237-45.
- Hooper, L. V., and Gordon, J. I. (2001) Glycans as legislators of host-microbial interactions: spanning the spectrum from symbiosis to pathogenicity. *Glycobiology* **11**, 1R-10R.

- Hu, K. J., and Leung, P. C. (2006) Complete, precise, and innocuous loss of multiple introns in the currently intronless, active cathepsin L-like genes, and inference from this event. *Mol Phylogenet Evol* **38**, 685-96.
- Hu, Q. D., Ang, B. T., Karsak, M., Hu, W. P., Cui, X. Y., Duka, T., Takeda, Y., Chia, W., Sankar, N., Ng, Y. K., Ling, E. A., Maciag, T., Small, D., Trifonova, R., Kopan, R., Okano, H., Nakafuku, M., Chiba, S., Hirai, H., Aster, J. C., Schachner, M., Pallen, C. J., Watanabe, K., and Xiao, Z. C. (2003) F3/contactin acts as a functional ligand for Notch during oligodendrocyte maturation. *Cell* **115**, 163-75.
- Huh, M. S., Smid, J. K., and Rudnicki, M. A. (2005) Muscle function and dysfunction in health and disease. *Birth Defects Res C Embryo Today* **75**, 180-92.
- Irvine, K. D. (1999) Fringe, Notch, and making developmental boundaries. *Curr Opin Genet Dev* **9**, 434-41.
- Irvine, K. D., and Wieschaus, E. (1994) fringe, a Boundary-specific signaling molecule, mediates interactions between dorsal and ventral cells during *Drosophila* wing development. *Cell* **79**, 595-606.
- Ishikawa, H. O., Higashi, S., Ayukawa, T., Sasamura, T., Kitagawa, M., Harigaya, K., Aoki, K., Ishida, N., Sanai, Y., and Matsuno, K. (2005) Notch deficiency implicated in the pathogenesis of congenital disorder of glycosylation IIc. *Proc Natl Acad Sci U S A* **102**, 18532-7.
- Iwahori, A., Fraidenraich, D., and Basilico, C. (2004) A conserved enhancer element that drives FGF4 gene expression in the embryonic myotomes is synergistically activated by GATA and bHLH proteins. *Dev Biol* **270**, 525-37.
- Jaeken, J. (2003) Komrower Lecture. Congenital disorders of glycosylation (CDG) it's all in it! *J Inherit Metab Dis* **26**, 99-118.
- Jaeken, J., and Matthijs, G. (2001) Congenital disorders of glycosylation. *Annu Rev Genomics Hum Genet* **2**, 129-51.
- Jarriault, S., Brou, C., Logeat, F., Schroeter, E. H., Kopan, R., and Israel, A. (1995) Signalling downstream of activated mammalian Notch. *Nature* **377**, 355-8.
- Javaud, C., Dupuy, F., Maftah, A., Julien, R., and Petit, J. M. (2003) The fucosyltransferase gene family: an amazing summary of the underlying mechanisms of gene evolution. *Genetica* **118**, 157-70.
- Javaud, C., Dupuy, F., Maftah, A., Michalski, J. C., Oriol, R., Petit, J. M., and Julien, R. (2000) Ancestral exonic organization of FUT8, the gene encoding the alpha6-fucosyltransferase, reveals successive peptide domains which suggest a particular

- three-dimensional core structure for the alpha6-fucosyltransferase family. *Mol Biol Evol* **17**, 1661-72.
- Jego, P., Joly, J., and Boisseau, C. (1980) [Amphibian jelly envelopes (proteins secreted by the oviduct) and their role in fertilization]. *Reprod Nutr Dev* **20**, 557-67.
- Johnston, S. H., Rauskolb, C., Wilson, R., Prabhakaran, B., Irvine, K. D., and Vogt, T. F. (1997) A family of mammalian Fringe genes implicated in boundary determination and the Notch pathway. *Development* **124**, 2245-54.
- Jordon, D., Jagirdar, J., and Kaneko, M. (1989) Blood group antigens, Lewisx and Lewisy in the diagnostic discrimination of malignant mesothelioma versus adenocarcinoma. *Am J Pathol* **135**, 931-7.
- Julien, S., Lagadec, C., Krzewinski-Recchi, M. A., Courtand, G., Le Bourhis, X., and Delannoy, P. (2005) Stable expression of sialyl-Tn antigen in T47-D cells induces a decrease of cell adhesion and an increase of cell migration. *Breast Cancer Res Treat* **90**, 77-84.
- Jurado, L. A., Coloma, A., and Cruces, J. (1999) Identification of a human homolog of the *Drosophila* rotated abdomen gene (POMT1) encoding a putative protein O-mannosyltransferase, and assignment to human chromosome 9q34.1. *Genomics* **58**, 171-80.
- Kabat, E. A., and McDuffie, F. C. (1956) A comparative study of methods used for analysis of specific precipitates in quantitative immunochemistry. *J Immunol* **77**, 193-7.
- Kablar, B., Krastel, K., Tajbakhsh, S., and Rudnicki, M. A. (2003) Myf5 and MyoD activation define independent myogenic compartments during embryonic development. *Dev Biol* **258**, 307-18.
- Kalchauer, C., Cinnamon, Y., and Kahane, N. (1999) Myotome formation: a multistage process. *Cell Tissue Res* **296**, 161-73.
- Kannagi, R., Shigeta, K., Hirashima, K., and Takada, A. (1989) [Clinical application of cancer-associated carbohydrate antigens in the diagnosis and therapy of human cancers]. *Gan To Kagaku Ryoho* **16**, 1157-66.
- Kao, Y. H., Lee, G. F., Wang, Y., Starovasnik, M. A., Kelley, R. F., Spellman, M. W., and Lerner, L. (1999) The effect of O-fucosylation on the first EGF-like domain from human blood coagulation factor VII. *Biochemistry* **38**, 7097-110.
- Kawano, Y., and Kypta, R. (2003) Secreted antagonists of the Wnt signalling pathway. *J Cell Sci* **116**, 2627-34.
- Kenny-Mobbs, T., and Thorogood, P. (1987) Autonomy of differentiation in avian branchial somites and the influence of adjacent tissues. *Development* **100**, 449-62.

- Kentzer, E. J., Buko, A., Menon, G., and Sarin, V. K. (1990) Carbohydrate composition and presence of a fucose-protein linkage in recombinant human pro-urokinase. *Biochem Biophys Res Commun* **171**, 401-6.
- Kidd, S., Lieber, T., and Young, M. W. (1998) Ligand-induced cleavage and regulation of nuclear entry of Notch in *Drosophila melanogaster* embryos. *Genes Dev* **12**, 3728-40.
- Kimble, J., and Simpson, P. (1997) The LIN-12/Notch signaling pathway and its regulation. *Annu Rev Cell Dev Biol* **13**, 333-61.
- Kitazume, S., Tachida, Y., Oka, R., Shirotani, K., Saido, T. C., and Hashimoto, Y. (2001) Alzheimer's beta-secretase, beta-site amyloid precursor protein-cleaving enzyme, is responsible for cleavage secretion of a Golgi-resident sialyltransferase. *Proc Natl Acad Sci U S A* **98**, 13554-9.
- Kleene, R., and Berger, E. G. (1993) The molecular and cell biology of glycosyltransferases. *Biochim Biophys Acta* **1154**, 283-325.
- Kobata, A. (1984) [The sugar chains of gamma-glutamyltranspeptidase--changes induced by malignant transformation and their application for the diagnosis of hepatoma]. *Yakugaku Zasshi* **104**, 823-38.
- Kocamis, H., Gahr, S. A., Batelli, L., Hubbs, A. F., and Killefer, J. (2002) IGF-I, IGF-II, and IGF-receptor-1 transcript and IGF-II protein expression in myostatin knockout mice tissues. *Muscle Nerve* **26**, 55-63.
- Kolset, S. O., and Gallagher, J. T. (1990) Proteoglycans in haemopoietic cells. *Biochim Biophys Acta* **1032**, 191-211.
- Kopan, R., Schroeter, E. H., Weintraub, H., and Nye, J. S. (1996) Signal transduction by activated mNotch: importance of proteolytic processing and its regulation by the extracellular domain. *Proc Natl Acad Sci U S A* **93**, 1683-8.
- Kornfeld, R., and Kornfeld, S. (1985) Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* **54**, 631-64.
- Kostova, Z., and Wolf, D. H. (2003) For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection. *Embo J* **22**, 2309-17.
- Kozak, M. (1989) The scanning model for translation: an update. *J Cell Biol* **108**, 229-41.
- Kukuruzinska, M. A., and Lennon, K. (1998) Protein N-glycosylation: molecular genetics and functional significance. *Crit Rev Oral Biol Med* **9**, 415-48.
- Kukuruzinska, M. A., and Lennon-Hopkins, K. (1999) ALG gene expression and cell cycle progression. *Biochim Biophys Acta* **1426**, 359-72.
- Kundra, R., and Kornfeld, S. (1999) Asparagine-linked oligosaccharides protect Lamp-1 and

- Lamp-2 from intracellular proteolysis. *J Biol Chem* **274**, 31039-46.
- Lareau, L. F., Green, R. E., Bhatnagar, R. S., and Brenner, S. E. (2004) The evolving roles of alternative splicing. *Curr Opin Struct Biol* **14**, 273-82.
- Lassar, A. B., Buskin, J. N., Lockshon, D., Davis, R. L., Apone, S., Hauschka, S. D., and Weintraub, H. (1989) MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer. *Cell* **58**, 823-31.
- Lecointre, G., and Le Guyader, H. (2001) Classification phylogénétique du vivant. Ed. Belin, Paris, 543 pp.
- Leeb, B. F., Petera, P., and Neumann, K. (1996) [Results of a multicenter study of chondroitin sulfate (Condrosulf) use in arthroses of the finger, knee and hip joints]. *Wien Med Wochenschr* **146**, 609-14.
- Lefebvre, T., Ferreira, S., Dupont-Wallois, L., Bussiere, T., Dupire, M. J., Delacourte, A., Michalski, J. C., and Caillet-Boudin, M. L. (2003) Evidence of a balance between phosphorylation and O-GlcNAc glycosylation of Tau proteins--a role in nuclear localization. *Biochim Biophys Acta* **1619**, 167-76.
- Legrand, C., Thibert, V., Dubernard, V., Begault, B., and Lawler, J. (1992) Molecular requirements for the interaction of thrombospondin with thrombin-activated human platelets: modulation of platelet aggregation. *Blood* **79**, 1995-2003.
- Lehrman, M. A. (2001) Oligosaccharide-based information in endoplasmic reticulum quality control and other biological systems. *J Biol Chem* **276**, 8623-6.
- Lei, L., Xu, A., Panin, V. M., and Irvine, K. D. (2003) An O-fucose site in the ligand binding domain inhibits Notch activation. *Development* **130**, 6411-21.
- Lis, H., and Sharon, N. (1993) Protein glycosylation. Structural and functional aspects. *Eur J Biochem* **218**, 1-27.
- Lissemore, J. L., and Starmer, W. T. (1999) Phylogenetic analysis of vertebrate and invertebrate Delta/Serrate/LAG-2 (DSL) proteins. *Mol Phylogenet Evol* **11**, 308-19.
- Liu, C. K., Schmied, R., Scher, W., and Waxman, S. (1981) Alterations in glycosyltransferase levels in mouse erythroleukemia cells during erythrodifferentiation and cell growth. *Cancer Res* **41**, 790-4.
- Logeat, F., Bessia, C., Brou, C., LeBail, O., Jarriault, S., Seidah, N. G., and Israel, A. (1998) The Notch1 receptor is cleaved constitutively by a furin-like convertase. *Proc Natl Acad Sci U S A* **95**, 8108-12.
- Loriol, C., Dupuy, F., Rampal, R., Dlugosz, M. A., Haltiwanger, R. S., Maftah, A., and

- Germot, A. (2006) Molecular evolution of protein O-fucosyltransferase genes and splice variants. *Glycobiology* **16**, 736-747.
- Lowe, J. B., and Marth, J. D. (2003) A genetic approach to Mammalian glycan function. *Annu Rev Biochem* **72**, 643-91.
- Lubas, W. A., Frank, D. W., Krause, M., and Hanover, J. A. (1997) O-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. *J Biol Chem* **272**, 9316-24.
- Lubke, T., Marquardt, T., Etzioni, A., Hartmann, E., von Figura, K., and Korner, C. (2001) Complementation cloning identifies CDG-IIc, a new type of congenital disorders of glycosylation, as a GDP-fucose transporter deficiency. *Nat Genet* **28**, 73-6.
- Lubke, T., Marquardt, T., von Figura, K., and Korner, C. (1999) A new type of carbohydrate-deficient glycoprotein syndrome due to a decreased import of GDP-fucose into the golgi. *J Biol Chem* **274**, 25986-9.
- Luo, D., Renault, V. M., and Rando, T. A. (2005) The regulation of Notch signaling in muscle stem cell activation and postnatal myogenesis. *Semin Cell Dev Biol* **16**, 612-22.
- Luo, Y., and Haltiwanger, R. S. (2005) O-fucosylation of notch occurs in the endoplasmic reticulum. *J Biol Chem* **280**, 11289-94.
- Luo, Y., Koles, K., Vorndam, W., Haltiwanger, R. S., and Panin, V. M. (2006) Protein O-fucosyltransferase 2 adds O-fucose to thrombospondin type 1 repeats. *J Biol Chem* **281**, 9393-9.
- Luo, Y., Nita-Lazar, A., and Haltiwanger, R. S. (2006) Two distinct pathways for O-fucosylation of epidermal growth factor-like or thrombospondin type 1 repeats. *J Biol Chem*.
- Ma, Y. Q., and Geng, J. G. (2000) Heparan sulfate-like proteoglycans mediate adhesion of human malignant melanoma A375 cells to P-selectin under flow. *J Immunol* **165**, 558-65.
- Mancias, J. D., and Goldberg, J. (2005) Exiting the endoplasmic reticulum. *Traffic* **6**, 278-85.
- Marquardt, T., and Denecke, J. (2003) Congenital disorders of glycosylation: review of their molecular bases, clinical presentations and specific therapies. *Eur J Pediatr* **162**, 359-79.
- Martin-Blanco, E., and Garcia-Bellido, A. (1996) Mutations in the rotated abdomen locus affect muscle development and reveal an intrinsic asymmetry in *Drosophila*. *Proc Natl Acad Sci U S A* **93**, 6048-52.
- Martinez-Duncker, I., Mollicone, R., Candelier, J. J., Breton, C., and Oriol, R. (2003) A new

- superfamily of protein-O-fucosyltransferases, alpha2-fucosyltransferases, and alpha6-fucosyltransferases: phylogeny and identification of conserved peptide motifs. *Glycobiology* **13**, 1C-5C.
- Meanwell, N. A., and Krystal, M. (2000) Respiratory syncytial virus: recent progress towards the discovery of effective prophylactic and therapeutic agents. *Drug Discov Today* **5**, 241-252.
- Menzel, O., Vellai, T., Takacs-Vellai, K., Reymond, A., Mueller, F., Antonarakis, S. E., and Guipponi, M. (2004) The *Caenorhabditis elegans* ortholog of C21orf80, a potential new protein O-fucosyltransferase, is required for normal development. *Genomics* **84**, 320-30.
- Mer, G., Hietter, H., and Lefevre, J. F. (1996) Stabilization of proteins by glycosylation examined by NMR analysis of a fucosylated proteinase inhibitor. *Nat Struct Biol* **3**, 45-53.
- Mercier, D., Wierinckx, A., Oulmouden, A., Gallet, P. F., Palcic, M. M., Harduin-Lepers, A., Delannoy, P., Petit, J. M., Leveziel, H., and Julien, R. (1999) Molecular cloning, expression and exon/intron organization of the bovine beta-galactoside alpha2,6-sialyltransferase gene. *Glycobiology* **9**, 851-63.
- Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J. W., and Marth, J. D. (1994) Complex asparagine-linked oligosaccharides are required for morphogenic events during post-implantation development. *Embo J* **13**, 2056-65.
- Miano, J. M. (2003) Serum response factor: toggling between disparate programs of gene expression. *J Mol Cell Cardiol* **35**, 577-93.
- Miller, J. R. (2002) The Wnts. *Genome Biol* **3**, REVIEWS3001.
- Molkentin, J. D., Black, B. L., Martin, J. F., and Olson, E. N. (1995) Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* **83**, 1125-36.
- Moloney, D. J., and Haltiwanger, R. S. (1999) The O-linked fucose glycosylation pathway: identification and characterization of a uridine diphosphoglucose: fucose-beta1,3-glucosyltransferase activity from Chinese hamster ovary cells. *Glycobiology* **9**, 679-87.
- Moloney, D. J., Lin, A. I., and Haltiwanger, R. S. (1997) The O-linked fucose glycosylation pathway. Evidence for protein-specific elongation of o-linked fucose in Chinese hamster ovary cells. *J Biol Chem* **272**, 19046-50.
- Moloney, D. J., Panin, V. M., Johnston, S. H., Chen, J., Shao, L., Wilson, R., Wang, Y., Stanley, P., Irvine, K. D., Haltiwanger, R. S., and Vogt, T. F. (2000) Fringe is a

- glycosyltransferase that modifies Notch. *Nature* **406**, 369-75.
- Moloney, D. J., Shair, L. H., Lu, F. M., Xia, J., Locke, R., Matta, K. L., and Haltiwanger, R. S. (2000) Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *J Biol Chem* **275**, 9604-11.
- Monnat, J., Neuhaus, E. M., Pop, M. S., Ferrari, D. M., Kramer, B., and Soldati, T. (2000) Identification of a novel saturable endoplasmic reticulum localization mechanism mediated by the C-terminus of a Dictyostelium protein disulfide isomerase. *Mol Biol Cell* **11**, 3469-84.
- Morisaki, T., Sermsuvitayawong, K., Byun, S. H., Matsuda, Y., Hidaka, K., Morisaki, H., and Mukai, T. (1997) Mouse Mef2b gene: unique member of MEF2 gene family. *J Biochem (Tokyo)* **122**, 939-46.
- Mosher, D. F. (1990) Physiology of thrombospondin. *Annu Rev Med* **41**, 85-97.
- Mumm, J. S., and Kopan, R. (2000) Notch signaling: from the outside in. *Dev Biol* **228**, 151-65.
- Munro, S., and Freeman, M. (2000) The notch signalling regulator fringe acts in the Golgi apparatus and requires the glycosyltransferase signature motif DXD. *Curr Biol* **10**, 813-20.
- Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B., and et al. (1989) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**, 537-44.
- Musaro, A., McCullagh, K. J., Naya, F. J., Olson, E. N., and Rosenthal, N. (1999) IGF-1 induces skeletal myocyte hypertrophy through calcineurin in association with GATA-2 and NF-ATc1. *Nature* **400**, 581-5.
- Naidu, P. S., Ludolph, D. C., To, R. Q., Hinterberger, T. J., and Konieczny, S. F. (1995) Myogenin and MEF2 function synergistically to activate the MRF4 promoter during myogenesis. *Mol Cell Biol* **15**, 2707-18.
- Naim, H. Y., and Lentze, M. J. (1992) Impact of O-glycosylation on the function of human intestinal lactase-phlorizin hydrolase. Characterization of glycoforms varying in enzyme activity and localization of O-glycoside addition. *J Biol Chem* **267**, 25494-504.
- Nakakura, N., Hietter, H., Van Dorsselaer, A., and Luu, B. (1992) Isolation and structural determination of three peptides from the insect *Locusta migratoria*. Identification of a

- deoxyhexose-linked peptide. *Eur J Biochem* **204**, 147-53.
- Nishimura, H., Kawabata, S., Kisiel, W., Hase, S., Ikenaka, T., Takao, T., Shimonishi, Y., and Iwanaga, S. (1989) Identification of a disaccharide (Xyl-Glc) and a trisaccharide (Xyl₂-Glc) O-glycosidically linked to a serine residue in the first epidermal growth factor-like domain of human factors VII and IX and protein Z and bovine protein Z. *J Biol Chem* **264**, 20320-5.
- Nishimura, H., Takao, T., Hase, S., Shimonishi, Y., and Iwanaga, S. (1992) Human factor IX has a tetrasaccharide O-glycosidically linked to serine 61 through the fucose residue. *J Biol Chem* **267**, 17520-5.
- Nishimura, H., Yamashita, S., Zeng, Z., Walz, D. A., and Iwanaga, S. (1992) Evidence for the existence of O-linked sugar chains consisting of glucose and xylose in bovine thrombospondin. *J Biochem (Tokyo)* **111**, 460-4.
- Nofziger, D., Miyamoto, A., Lyons, K. M., and Weinmaster, G. (1999) Notch signaling imposes two distinct blocks in the differentiation of C2C12 myoblasts. *Development* **126**, 1689-702.
- Oguri, S., Minowa, M. T., Ihara, Y., Taniguchi, N., Ikenaga, H., and Takeuchi, M. (1997) Purification and characterization of UDP-N-acetylglucosamine: alpha1,3-D-mannoside beta1,4-N-acetylglucosaminyltransferase (N-acetylglucosaminyltransferase-IV) from bovine small intestine. *J Biol Chem* **272**, 22721-7.
- Oh-eda, M., Hasegawa, M., Hattori, K., Kuboniwa, H., Kojima, T., Orita, T., Tomonou, K., Yamazaki, T., and Ochi, N. (1990) O-linked sugar chain of human granulocyte colony-stimulating factor protects it against polymerization and denaturation allowing it to retain its biological activity. *J Biol Chem* **265**, 11432-5.
- Okajima, T., and Irvine, K. D. (2002) Regulation of notch signaling by o-linked fucose. *Cell* **111**, 893-904.
- Okajima, T., Xu, A., and Irvine, K. D. (2003) Modulation of notch-ligand binding by protein O-fucosyltransferase 1 and fringe. *J Biol Chem* **278**, 42340-5.
- Okajima, T., Xu, A., Lei, L., and Irvine, K. D. (2005) Chaperone activity of protein O-fucosyltransferase 1 promotes notch receptor folding. *Science* **307**, 1599-603.
- Olczak, M., and Watorek, W. (2000) Structural analysis of N-glycans from yellow lupin (*Lupinus luteus*) seed diphosphonucleotide phosphatase/phosphodiesterase. *Biochim Biophys Acta* **1523**, 236-45.
- Oliver, J. D., van der Wal, F. J., Bulleid, N. J., and High, S. (1997) Interaction of the thiol-

- dependent reductase ERp57 with nascent glycoproteins. *Science* **275**, 86-8.
- Olson, E. N. (1993) Signal transduction pathways that regulate skeletal muscle gene expression. *Mol Endocrinol* **7**, 1369-78.
- Olson, E. N., and Klein, W. H. (1994) bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. *Genes Dev* **8**, 1-8.
- Ordahl, C. P., and Le Douarin, N. M. (1992) Two myogenic lineages within the developing somite. *Development* **114**, 339-53.
- Oriol, R., Mollicone, R., Cailleau, A., Balanzino, L., and Breton, C. (1999) Divergent evolution of fucosyltransferase genes from vertebrates, invertebrates, and bacteria. *Glycobiology* **9**, 323-34.
- Oulmouden, A., Wierinckx, A., Petit, J. M., Costache, M., Palcic, M. M., Mollicone, R., Oriol, R., and Julien, R. (1997) Molecular cloning and expression of a bovine alpha(1,3)-fucosyltransferase gene homologous to a putative ancestor gene of the human FUT3-FUT5-FUT6 cluster. *J Biol Chem* **272**, 8764-73.
- Pagny, S., Cabanes-Macheteau, M., Gillikin, J. W., Leborgne-Castel, N., Lerouge, P., Boston, R. S., Faye, L., and Gomord, V. (2000) Protein recycling from the Golgi apparatus to the endoplasmic reticulum in plants and its minor contribution to calreticulin retention. *Plant Cell* **12**, 739-56.
- Pan, Y. T., and Elbein, A. D. (1990) Control of N-linked oligosaccharide synthesis: cellular levels of dolichyl phosphate are not the only regulatory factor. *Biochemistry* **29**, 8077-84.
- Panin, V. M., Shao, L., Lei, L., Moloney, D. J., Irvine, K. D., and Haltiwanger, R. S. (2002) Notch ligands are substrates for protein O-fucosyltransferase-1 and Fringe. *J Biol Chem* **277**, 29945-52.
- Papy-Garcia, D., Barbosa, I., Duchesnay, A., Saadi, S., Caruelle, J. P., Barritault, D., and Martelly, I. (2002) Glycosaminoglycan mimetics (RGTA) modulate adult skeletal muscle satellite cell proliferation in vitro. *J Biomed Mater Res* **62**, 46-55.
- Pelham, H. R. (1996) The dynamic organisation of the secretory pathway. *Cell Struct Funct* **21**, 413-9.
- Perez-Vilar, J., and Hill, R. L. (1999) The structure and assembly of secreted mucins. *J Biol Chem* **274**, 31751-4.
- Peterson, R. E., and McClay, D. R. (2005) A Fringe-modified Notch signal affects specification of mesoderm and endoderm in the sea urchin embryo. *Dev Biol* **282**, 126-37.

- Petit, D., Maftah, A., Julien, R., and Petit, JM. (2006) En bloc duplications, mutation rates and densities of amino acid changes clarify the evolution of vertebrate α 1,3/4 fucosyltransferases. *J. Mol. Evol.*, in press.
- Petitou, M., Lormeau, J. C., and Choay, J. (1988) Interaction of heparin and antithrombin III. The role of O-sulfate groups. *Eur J Biochem* **176**, 637-40.
- Piccolo, S., Agius, E., Leyns, L., Bhattacharyya, S., Grunz, H., Bouwmeester, T., and De Robertis, E. M. (1999) The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature* **397**, 707-10.
- Pirskanen, A., Kiefer, J. C., and Hauschka, S. D. (2000) IGFs, insulin, Shh, bFGF, and TGF-beta1 interact synergistically to promote somite myogenesis in vitro. *Dev Biol* **224**, 189-203.
- Pollevick, G. D., Di Noia, J. M., Salto, M. L., Lima, C., Leguizamon, M. S., de Lederkremer, R. M., and Frasch, A. C. (2000) Trypanosoma cruzi surface mucins with exposed variant epitopes. *J Biol Chem* **275**, 27671-80.
- Pourquie, O., Fan, C. M., Coltey, M., Hirsinger, E., Watanabe, Y., Breant, C., Francis-West, P., Brickell, P., Tessier-Lavigne, M., and Le Douarin, N. M. (1996) Lateral and axial signals involved in avian somite patterning: a role for BMP4. *Cell* **84**, 461-71.
- Rabbani, S. A., Mazar, A. P., Bernier, S. M., Haq, M., Bolivar, I., Henkin, J., and Goltzman, D. (1992) Structural requirements for the growth factor activity of the amino-terminal domain of urokinase. *J Biol Chem* **267**, 14151-6.
- Rademacher, T. W., Parekh, R. B., and Dwek, R. A. (1988) Glycobiology. *Annu Rev Biochem* **57**, 785-838.
- Rampal, R., Arboleda-Velasquez, J. F., Nita-Lazar, A., Kosik, K. S., and Haltiwanger, R. S. (2005) Highly conserved O-fucose sites have distinct effects on Notch1 function. *J Biol Chem* **280**, 32133-40.
- Rampal, R., Li, A. S., Moloney, D. J., Georgiou, S. A., Luther, K. B., Nita-Lazar, A., and Haltiwanger, R. S. (2005) Lunatic fringe, manic fringe, and radical fringe recognize similar specificity determinants in O-fucosylated epidermal growth factor-like repeats. *J Biol Chem* **280**, 42454-63.
- Rapoport, T. A., Matlack, K. E., Plath, K., Misselwitz, B., and Staech, O. (1999) Posttranslational protein translocation across the membrane of the endoplasmic reticulum. *Biol Chem* **380**, 1143-50.
- Rastan, S. (1994) X chromosome inactivation and the Xist gene. *Curr Opin Genet Dev* **4**, 292-7.

- Rawls, A., and Olson, E. N. (1997) MyoD meets its maker. *Cell* **89**, 5-8.
- Rhodes, S. J., and Konieczny, S. F. (1989) Identification of MRF4: a new member of the muscle regulatory factor gene family. *Genes Dev* **3**, 2050-61.
- Ridgeway, A. G., Petropoulos, H., Wilton, S., and Skerjanc, I. S. (2000) Wnt signaling regulates the function of MyoD and myogenin. *J Biol Chem* **275**, 32398-405.
- Rijsewijk, F., Schuermann, M., Wagenaar, E., Parren, P., Weigel, D., and Nusse, R. (1987) The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell* **50**, 649-57.
- Roberts, D. D. (1996) Regulation of tumor growth and metastasis by thrombospondin-1. *Faseb J* **10**, 1183-91.
- Rogerson, P. J., Jamali, M., and Skerjanc, I. S. (2002) The C-terminus of myogenin, but not MyoD, targets upregulation of MEF2C expression. *FEBS Lett* **524**, 134-8.
- Roos, C., Kolmer, M., Mattila, P., and Renkonen, R. (2002) Composition of Drosophila melanogaster proteome involved in fucosylated glycan metabolism. *J Biol Chem* **277**, 3168-75.
- Roseman, S. (1970) The synthesis of complex carbohydrates by multiglycosyltransferase systems and their potential function in intercellular adhesion. *Chem Phys Lipids* **5**, 270-97.
- Roth, J. (2002) Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. *Chem Rev* **102**, 285-303.
- Rudnicki, M. A., Braun, T., Hinuma, S., and Jaenisch, R. (1992) Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell* **71**, 383-90.
- Runte, M., Kroisel, P. M., Gillessen-Kaesbach, G., Varon, R., Horn, D., Cohen, M. Y., Wagstaff, J., Horsthemke, B., and Buiting, K. (2004) SNURF-SNRPN and UBE3A transcript levels in patients with Angelman syndrome. *Hum Genet* **114**, 553-61.
- Sabourin, L. A., and Rudnicki, M. A. (2000) The molecular regulation of myogenesis. *Clin Genet* **57**, 16-25.
- Sakamoto, K., Ohara, O., Takagi, M., Takeda, S., and Katsube, K. (2002) Intracellular cell-autonomous association of Notch and its ligands: a novel mechanism of Notch signal modification. *Dev Biol* **241**, 313-26.
- Sakamoto, S., Watanabe, T., Tokumaru, T., Takagi, H., Nakazato, H., and Lloyd, K. O. (1989) Expression of Lewis^a, Lewis^b, Lewis^x, Lewis^y, sialyl-Lewis^a, and sialyl-

- Lewisx blood group antigens in human gastric carcinoma and in normal gastric tissue. *Cancer Res* **49**, 745-52.
- Sakamoto, Y., Taguchi, T., Honke, K., Korekane, H., Watanabe, H., Tano, Y., Dohmae, N., Takio, K., Horii, A., and Taniguchi, N. (2000) Molecular cloning and expression of cDNA encoding chicken UDP-N-acetyl-D-glucosamine (GlcNAc) GlcNAc β 1-6(GlcNAc β 1-2)- man α 1-R[GlcNAc to man] β 1,4N-acetylglucosaminyltransferase VI. *J Biol Chem* **275**, 36029-34.
- Sanchez-Irizarry, C., Carpenter, A. C., Weng, A. P., Pear, W. S., Aster, J. C., and Blacklow, S. C. (2004) Notch subunit heterodimerization and prevention of ligand-independent proteolytic activation depend, respectively, on a novel domain and the LNR repeats. *Mol Cell Biol* **24**, 9265-73.
- Sanes, J. R. (2003) The basement membrane/basal lamina of skeletal muscle. *J Biol Chem* **278**, 12601-4.
- Santerre, R. F., Bales, K. R., Janney, M. J., Hannon, K., Fisher, L. F., Bailey, C. S., Morris, J., Ivarie, R., and Smith, C. K., 2nd (1993) Expression of bovine myf5 induces ectopic skeletal muscle formation in transgenic mice. *Mol Cell Biol* **13**, 6044-51.
- Sasaki, T., Yamada, H., Matsumura, K., Shimizu, T., Kobata, A., and Endo, T. (1998) Detection of O-mannosyl glycans in rabbit skeletal muscle alpha-dystroglycan. *Biochim Biophys Acta* **1425**, 599-606.
- Sasamura, T., Sasaki, N., Miyashita, F., Nakao, S., Ishikawa, H. O., Ito, M., Kitagawa, M., Harigaya, K., Spana, E., Bilder, D., Perrimon, N., and Matsuno, K. (2003) neurotic, a novel maternal neurogenic gene, encodes an O-fucosyltransferase that is essential for Notch-Delta interactions. *Development* **130**, 4785-95.
- Saunier, K., Barreaud, J. P., Eggen, A., Oriol, R., Leveziel, H., Julien, R., and Petit, J. M. (2001) Organization of the bovine alpha 2-fucosyltransferase gene cluster suggests that the Sec1 gene might have been shaped through a nonautonomous L1-retrotransposition event within the same locus. *Mol Biol Evol* **18**, 2083-91.
- Schiffer, S. G., Foley, S., Kaffashan, A., Hronowski, X., Zichittella, A. E., Yeo, C. Y., Miatkowski, K., Adkins, H. B., Damon, B., Whitman, M., Salomon, D., Sanicola, M., and Williams, K. P. (2001) Fucosylation of Cripto is required for its ability to facilitate nodal signaling. *J Biol Chem* **276**, 37769-78.
- Schonherr, E., Hausser, H., Beavan, L., and Kresse, H. (1995) Decorin-type I collagen interaction. Presence of separate core protein-binding domains. *J Biol Chem* **270**, 8877-83.

- Schroeter, E. H., Kisslinger, J. A., and Kopan, R. (1998) Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain. *Nature* **393**, 382-6.
- Schweisguth, F., Gho, M., and Lecourtois, M. (1996) Control of cell fate choices by lateral signaling in the adult peripheral nervous system of *Drosophila melanogaster*. *Dev Genet* **18**, 28-39.
- Schwientek, T., Bennett, E. P., Flores, C., Thacker, J., Hollmann, M., Reis, C. A., Behrens, J., Mandel, U., Keck, B., Schafer, M. A., Haselmann, K., Zubarev, R., Roepstorff, P., Burchell, J. M., Taylor-Papadimitriou, J., Hollingsworth, M. A., and Clausen, H. (2002) Functional conservation of subfamilies of putative UDP-N-acetylgalactosamine:polypeptide N-acetylgalactosaminyltransferases in *Drosophila*, *Caenorhabditis elegans*, and mammals. One subfamily composed of I(2)35Aa is essential in *Drosophila*. *J Biol Chem* **277**, 22623-38.
- Scott, D. B., Blanpied, T. A., and Ehlers, M. D. (2003) Coordinated PKA and PKC phosphorylation suppresses RXR-mediated ER retention and regulates the surface delivery of NMDA receptors. *Neuropharmacology* **45**, 755-67.
- Semsarian, C., Sutrave, P., Richmond, D. R., and Graham, R. M. (1999) Insulin-like growth factor (IGF-I) induces myotube hypertrophy associated with an increase in anaerobic glycolysis in a clonal skeletal-muscle cell model. *Biochem J* **339** (Pt 2), 443-51.
- Shafi, R., Iyer, S. P., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., and Marth, J. D. (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc Natl Acad Sci U S A* **97**, 5735-9.
- Shao, L., and Haltiwanger, R. S. (2003) O-fucose modifications of epidermal growth factor-like repeats and thrombospondin type 1 repeats: unusual modifications in unusual places. *Cell Mol Life Sci* **60**, 241-50.
- Shao, L., Luo, Y., Moloney, D. J., and Haltiwanger, R. (2002) O-glycosylation of EGF repeats: identification and initial characterization of a UDP-glucose: protein O-glucosyltransferase. *Glycobiology* **12**, 763-70.
- Shao, L., Moloney, D. J., and Haltiwanger, R. (2003) Fringe modifies O-fucose on mouse Notch1 at epidermal growth factor-like repeats within the ligand-binding site and the Abruptex region. *J Biol Chem* **278**, 7775-82.
- Shi, S., and Stanley, P. (2003) Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways. *Proc Natl Acad Sci U S A* **100**, 5234-9.
- Shi, S., and Stanley, P. (2006) Evolutionary origins of Notch signaling in early development.

- Cell Cycle* **5**, 274-8.
- Shi, Y., and Massague, J. (2003) Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* **113**, 685-700.
- Shimizu, K., Chiba, S., Saito, T., Kumano, K., Takahashi, T., and Hirai, H. (2001) Manic fringe and lunatic fringe modify different sites of the Notch2 extracellular region, resulting in different signaling modulation. *J Biol Chem* **276**, 25753-8.
- Shore, P., and Sharrocks, A. D. (1995) The MADS-box family of transcription factors. *Eur J Biochem* **229**, 1-13.
- Silberstein, S., and Gilmore, R. (1996) Biochemistry, molecular biology, and genetics of the oligosaccharyltransferase. *Faseb J* **10**, 849-58.
- Smalheiser, N. R., Haslam, S. M., Sutton-Smith, M., Morris, H. R., and Dell, A. (1998) Structural analysis of sequences O-linked to mannose reveals a novel Lewis X structure in cranin (dystroglycan) purified from sheep brain. *J Biol Chem* **273**, 23698-703.
- Spiro, R. G. (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43R-56R.
- Srikantan, V., Zou, Z., Petrovics, G., Xu, L., Augustus, M., Davis, L., Livezey, J. R., Connell, T., Sesterhenn, I. A., Yoshino, K., Buzard, G. S., Mostofi, F. K., McLeod, D. G., Moul, J. W., and Srivastava, S. (2000) PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. *Proc Natl Acad Sci U S A* **97**, 12216-21.
- Steele, R. E. (2002) Developmental signaling in Hydra: what does it take to build a "simple" animal? *Dev Biol* **248**, 199-219.
- Stern, H. M., Lin-Jones, J., and Hauschka, S. D. (1997) Synergistic interactions between bFGF and a TGF-beta family member may mediate myogenic signals from the neural tube. *Development* **124**, 3511-23.
- Strahl-Bolsinger, S., Gentsch, M., and Tanner, W. (1999) Protein O-mannosylation. *Biochim Biophys Acta* **1426**, 297-307.
- Strous, G. J., and Dekker, J. (1992) Mucin-type glycoproteins. *Crit Rev Biochem Mol Biol* **27**, 57-92.
- Sturla, L., Puglielli, L., Tonetti, M., Berninsone, P., Hirschberg, C. B., De Flora, A., and Etzioni, A. (2001) Impairment of the Golgi GDP-L-fucose transport and unresponsiveness to fucose replacement therapy in LAD II patients. *Pediatr Res* **49**, 537-42.
- Tamaki, T., Akatsuka, A., Okada, Y., Matsuzaki, Y., Okano, H., and Kimura, M. (2003)

- Growth and differentiation potential of main- and side-population cells derived from murine skeletal muscle. *Exp Cell Res* **291**, 83-90.
- Tanner, W., and Lehle, L. (1987) Protein glycosylation in yeast. *Biochim Biophys Acta* **906**, 81-99.
- Templeton, T. J., Iyer, L. M., Anantharaman, V., Enomoto, S., Abrahante, J. E., Subramanian, G. M., Hoffman, S. L., Abrahamsen, M. S., and Aravind, L. (2004) Comparative analysis of apicomplexa and genomic diversity in eukaryotes. *Genome Res* **14**, 1686-95.
- Ten Hagen, K. G., Fritz, T. A., and Tabak, L. A. (2003) All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology* **13**, 1R-16R.
- Torres, C. R., and Hart, G. W. (1984) Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *J Biol Chem* **259**, 3308-17.
- Tozeren, A., Kleinman, H. K., Grant, D. S., Morales, D., Mercurio, A. M., and Byers, S. W. (1995) E-selectin-mediated dynamic interactions of breast- and colon-cancer cells with endothelial-cell monolayers. *Int J Cancer* **60**, 426-31.
- Valdez, M. R., Richardson, J. A., Klein, W. H., and Olson, E. N. (2000) Failure of Myf5 to support myogenic differentiation without myogenin, MyoD, and MRF4. *Dev Biol* **219**, 287-98.
- Van den Steen, P., Rudd, P. M., Dwek, R. A., and Opdenakker, G. (1998) Concepts and principles of O-linked glycosylation. *Crit Rev Biochem Mol Biol* **33**, 151-208.
- Van Reeuwijk, J., Janssen, M., van den Elzen, C., Beltran-Valero de Bernabe, D., Sabatelli, P., Merlini, L., Boon, M., Scheffer, H., Brockington, M., Muntoni, F., Huynen, M. A., Verrips, A., Walsh, C. A., Barth, P. G., Brunner, H. G., and van Bokhoven, H. (2005) POMT2 mutations cause alpha-dystroglycan hypoglycosylation and Walker-Warburg syndrome. *J Med Genet* **42**, 907-12.
- Vandromme, M., Gauthier-Rouviere, C., Carnac, G., Lamb, N., and Fernandez, A. (1992) Serum response factor p67SRF is expressed and required during myogenic differentiation of both mouse C2 and rat L6 muscle cell lines. *J Cell Biol* **118**, 1489-500.
- Veach, Y. K., Martin, R. C., Mok, D. W., Malbeck, J., Vankova, R., and Mok, M. C. (2003) O-glucosylation of cis-zeatin in maize. Characterization of genes, enzymes, and endogenous cytokinins. *Plant Physiol* **131**, 1374-80.
- Velleman, S. G. (1999) The role of the extracellular matrix in skeletal muscle development.

- Poult Sci* **78**, 778-84.
- Velleman, S. G., Liu, X., Eggen, K. H., and Nestor, K. E. (1999) Developmental regulation of proteoglycan synthesis and decorin expression during turkey embryonic skeletal muscle formation. *Poult Sci* **78**, 1619-26.
- Venters, S. J., and Ordahl, C. P. (2002) Persistent myogenic capacity of the dermomyotome dorsomedial lip and restriction of myogenic competence. *Development* **129**, 3873-85.
- Venters, S. J., Thorsteinsdottir, S., and Duxson, M. J. (1999) Early development of the myotome in the mouse. *Dev Dyn* **216**, 219-32.
- Venuti, J. M., Morris, J. H., Vivian, J. L., Olson, E. N., and Klein, W. H. (1995) Myogenin is required for late but not early aspects of myogenesis during mouse development. *J Cell Biol* **128**, 563-76.
- Vidal y Plana, R. R., and Karzel, K. (1980) [Glucosamine: its value for the metabolism of articular cartilage. 1. Biochemistry of proteoglycans, studies on in-vitro cultures of embryonal mouse fibroblasts and bone germs]. *Fortschr Med* **98**, 557-62.
- Vijay, I. K. (1998) Developmental and hormonal regulation of protein N-glycosylation in the mammary gland. *J Mammary Gland Biol Neoplasia* **3**, 325-36.
- Villena, J., and Brandan, E. (2004) Dermatan sulfate exerts an enhanced growth factor response on skeletal muscle satellite cell proliferation and migration. *J Cell Physiol* **198**, 169-78.
- Vishwakarma, R. A., and Menon, A. K. (2005) Flip-flop of glycosylphosphatidylinositols (GPI's) across the ER. *Chem Commun (Camb)*, 453-5.
- Wakelam, M. J. (1985) The fusion of myoblasts. *Biochem J* **228**, 1-12.
- Wang, D. Z., Valdez, M. R., McAnally, J., Richardson, J., and Olson, E. N. (2001) The Mef2c gene is a direct transcriptional target of myogenic bHLH and MEF2 proteins during skeletal muscle development. *Development* **128**, 4623-33.
- Wang, Y., Lee, G. F., Kelley, R. F., and Spellman, M. W. (1996) Identification of a GDP-L-fucose:polypeptide fucosyltransferase and enzymatic addition of O-linked fucose to EGF domains. *Glycobiology* **6**, 837-42.
- Wang, Y., Shao, L., Shi, S., Harris, R. J., Spellman, M. W., Stanley, P., and Haltiwanger, R. S. (2001) Modification of epidermal growth factor-like repeats with O-fucose. Molecular cloning and expression of a novel GDP-fucose protein O-fucosyltransferase. *J Biol Chem* **276**, 40338-45.
- Wang, Y., and Spellman, M. W. (1998) Purification and characterization of a GDP-fucose:polypeptide fucosyltransferase from Chinese hamster ovary cells. *J Biol Chem*

- 273**, 8112-8.
- Wei, L., Zhou, W., Croissant, J. D., Johansen, F. E., Prywes, R., Balasubramanyam, A., and Schwartz, R. J. (1998) RhoA signaling via serum response factor plays an obligatory role in myogenic differentiation. *J Biol Chem* **273**, 30287-94.
- Weinmaster, G. (1997) The ins and outs of notch signaling. *Mol Cell Neurosci* **9**, 91-102.
- Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T. K., Turner, D., Rupp, R., Hollenberg, S., and et al. (1991) The myoD gene family: nodal point during specification of the muscle cell lineage. *Science* **251**, 761-6.
- Wells, L., Vosseller, K., and Hart, G. W. (2003) A role for N-acetylglucosamine as a nutrient sensor and mediator of insulin resistance. *Cell Mol Life Sci* **60**, 222-8.
- Wharton, K. A., Johansen, K. M., Xu, T., and Artavanis-Tsakonas, S. (1985) Nucleotide sequence from the neurogenic locus notch implies a gene product that shares homology with proteins containing EGF-like repeats. *Cell* **43**, 567-81.
- Wierinckx, A., Mercier, D., Oulmouden, A., Petit, J. M., and Julien, R. (1999) Complete genomic organization of futb encoding a bovine alpha 3-fucosyltransferase: exons in human orthologous genes emerged from ancestral intronic sequences. *Mol Biol Evol* **16**, 1535-47.
- Willer, T., Amselgruber, W., Deutzmann, R., and Strahl, S. (2002) Characterization of POMT2, a novel member of the PMT protein O-mannosyltransferase family specifically localized to the acrosome of mammalian spermatids. *Glycobiology* **12**, 771-83.
- Wilson-Rawls, J., Molkentin, J. D., Black, B. L., and Olson, E. N. (1999) Activated notch inhibits myogenic activity of the MADS-Box transcription factor myocyte enhancer factor 2C. *Mol Cell Biol* **19**, 2853-62.
- Wormald, M. R., and Dwek, R. A. (1999) Glycoproteins: glycan presentation and protein-fold stability. *Structure* **7**, R155-60.
- Wright, W. E., Sassoon, D. A., and Lin, V. K. (1989) Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell* **56**, 607-17.
- Yaffe, D. (1968) Retention of differentiation potentialities during prolonged cultivation of myogenic cells. *Proc Natl Acad Sci U S A* **61**, 477-83.
- Yaffe, D., and Saxel, O. (1977) Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature* **270**, 725-7.
- Yamauchi, M., Noyes, C., Kuboki, Y., and Mechanic, G. L. (1982) Collagen structural microheterogeneity and a possible role for glycosylated hydroxylysine in type I

- collagen. *Proc Natl Acad Sci U S A* **79**, 7684-8.
- Yan, Y. T., Liu, J. J., Luo, Y., E, C., Haltiwanger, R. S., Abate-Shen, C., and Shen, M. M. (2002) Dual roles of Cripto as a ligand and coreceptor in the nodal signaling pathway. *Mol Cell Biol* **22**, 4439-49.
- Yao, P. J., and Coleman, P. D. (1998) Reduction of O-linked N-acetylglucosamine-modified assembly protein-3 in Alzheimer's disease. *J Neurosci* **18**, 2399-411.
- Ye, Z., and Marth, J. D. (2004) N-glycan branching requirement in neuronal and postnatal viability. *Glycobiology* **14**, 547-58.
- Yoshida, A., Kobayashi, K., Manya, H., Taniguchi, K., Kano, H., Mizuno, M., Inazu, T., Mitsuhashi, H., Takahashi, S., Takeuchi, M., Herrmann, R., Straub, V., Talim, B., Voit, T., Topaloglu, H., Toda, T., and Endo, T. (2001) Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Dev Cell* **1**, 717-24.
- Young, H. E., Carrino, D. A., and Caplan, A. I. (1990) Change in synthesis of sulfated glycoconjugates during muscle development, maturation and aging in embryonic to senescent CBF-1 mouse. *Mech Ageing Dev* **53**, 179-93.
- Yu, Y. T., Breitbart, R. E., Smoot, L. B., Lee, Y., Mahdavi, V., and Nadal-Ginard, B. (1992) Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. *Genes Dev* **6**, 1783-98.
- Yuan, Y. P., Schultz, J., Mlodzik, M., and Bork, P. (1997) Secreted fringe-like signaling molecules may be glycosyltransferases. *Cell* **88**, 9-11.
- Zhang, J., and Xu, W. (2006) Recent advances in anti-influenza agents with neuraminidase as target. *Mini Rev Med Chem* **6**, 429-48.
- Zhang, N., Norton, C. R., and Gridley, T. (2002) Segmentation defects of Notch pathway mutants and absence of a synergistic phenotype in lunatic fringe/radical fringe double mutant mice. *Genesis* **33**, 21-8.
- Zhang, W., Betel, D., and Schachter, H. (2002) Cloning and expression of a novel UDP-GlcNAc:alpha-D-mannoside beta1,2-N-acetylglucosaminyltransferase homologous to UDP-GlcNAc:alpha-3-D-mannoside beta1,2-N-acetylglucosaminyltransferase I. *Biochem J* **361**, 153-62.
- Zhang, Y., Iwasaki, H., Wang, H., Kudo, T., Kalka, T. B., Hennes, T., Kubota, T., Cheng, L., Inaba, N., Gotoh, M., Togayachi, A., Guo, J., Hisatomi, H., Nakajima, K., Nishihara, S., Nakamura, M., Marth, J. D., and Narimatsu, H. (2003) Cloning and characterization of a new human UDP-N-acetyl-alpha-D-galactosamine:polypeptide

- N-acetylgalactosaminyltransferase, designated pp-GalNAc-T13, that is specifically expressed in neurons and synthesizes GalNAc alpha-serine/threonine antigen. *J Biol Chem* **278**, 573-84.
- Zimowska, M., Szczepankowska, D., Streminska, W., Papy, D., Tournaire, M. C., Gautron, J., Barritault, D., Moraczewski, J., and Martelly, I. (2001) Heparan sulfate mimetics modulate calpain activity during rat Soleus muscle regeneration. *J Cell Physiol* **188**, 178-87.
- Zorzano, A., Kaliman, P., Guma, A., and Palacin, M. (2003) Intracellular signals involved in the effects of insulin-like growth factors and neuregulins on myofibre formation. *Cell Signal* **15**, 141-9.

Les *O*-fucosyltransférases. Caractérisation des enzymes bovines et étude préliminaire du rôle de Pofut1 murine dans la différenciation de la cellule musculaire.

La *O*-fucosylation est l'ajout d'un fucose sur un résidu sérine ou thréonine compris dans deux types de domaines peptidiques, les EGF et les TSR. Cette modification post-traductionnelle dépend de deux enzymes, Pofut1 pour les EGF et Pofut2 pour les TSR. Nous avons étudié l'évolution des gènes *Pofut1* et *Pofut2*, et avons démontré que ces gènes présents en un seul exemplaire, existaient déjà chez l'ancêtre des Bilatériens voire des Métazoaires, probablement sous forme morcelée. Une situation originale existe pour *Pofut2*, retrouvé chez un groupe de Protozoaires, les Apicomplexés. La structure des deux gènes chez le bovin et leur expression tissulaire ont été établies. Nous trouvons une situation inédite puisqu'il existe pour chacun d'entre eux, cinq variants transcriptionnels dont un seul code l'enzyme active, différemment exprimés selon les tissus bovins. Les enzymes Pofut1 et Pofut2 actives seraient présentes dans tous les tissus analysés, à l'exception notable des muscles squelettiques de l'adulte où des formes atypiques sont présentes. Les autres variants transcriptionnels, la plupart tronqués, auraient un rôle dans la régulation du taux d'expression des gènes. Pofut1 est la première fucosyltransférase à avoir été identifiée comme résidente du réticulum endoplasmique. D'un point de vue fonctionnel, nous avons démontré chez le bovin que cette glycosyltransférase, porteuse de deux *N*-glycanes vraisemblablement de type oligomannosidique, était correctement repliée et donc douée d'une activité enzymatique, à condition que son premier site de *N*-glycosylation soit occupé. L'activité des récepteurs Notch et de leurs ligands présents à la surface de nombreuses cellules est modulée par l'état de *O*-fucosylation de leurs domaines EGF. Etant donnée l'implication de ces récepteurs dans la régulation de la myogenèse, la *O*-fucosylation doit y contribuer tout autant. Ainsi, nous avons débuté une étude des répercussions d'une modulation de l'expression de Pofut1 sur la différenciation de cellules primaires myoblastiques bovines et de cellules murines de la lignée C2C12. Nous montrons que la surexpression transitoire de l'enzyme Pofut1 murine retarde l'expression des facteurs de transcription myogéniques de la famille bHLH: Myf5, MyoD, Myogénine et MRF4. Ces résultats, encore préliminaires, ouvrent de nouvelles et passionnantes perspectives d'analyse de l'influence des *O*-fucosyltransférases au cours du processus de myogenèse.

Mots clés : *O*-fucosylation, organisation génomique, variants transcriptionnels, phylogénie, activité *O*-fucosyltransférase, myogenèse, bovin

The *O*-fucosyltransferases. Characterisation of the bovine enzymes and preliminary study of murine Pofut1 role in the differentiation of muscular cell.

O-fucosylation is the addition of a fucose on serine or threonine comprised in two types of peptidic domains, EGF and TSR. This post-translational modification depends on two enzymes. Pofut1 is responsible for *O*-fucosylation of EGF repeats, whereas Pofut2 adds fucose on TSR. We studied the evolution of *Pofut1* and *Pofut2* genes and demonstrated that these genes, present in a single copy, already existed in the ancestor of Bilaterians, or even Metazoans, probably in polyexonic form. An original situation exists for *Pofut2*, recovered in a group of Protozoans, the *Apicomplexa*. Structures of the two bovine genes and their tissular expression have been established. We find an original situation since it exists for each of them, five transcript variants of which only one encodes the active enzyme, differently expressed among bovine tissues. The Pofut1 and Pofut2 active enzymes would be present in all analyzed tissues, except for adult skeletal muscles where atypical forms are present. The other transcript variants, more or less truncated probably play a role in regulating the expression level of these genes. Pofut1 is the first fucosyltransferase to have been identified as an endoplasmic reticulum resident. Functionally, we demonstrated for the bovine that this glycosyltransferase, bearing two *N*-glycans probably of oligomannosidic type, was correctly folded and therefore possessed an enzymatic activity, provided its first site of *N*-glycosylation is occupied. The activity of the Notch receptors and their ligands, present at the surface of numerous cells, is modulated by the state of *O*-fucosylation of their EGF domains. Considering the implication of these receptors in the regulation of myogenesis, *O*-fucosylation must contribute all as much there. Thus, we started a study of the repercussions of a modulation of Pofut1 expression on the differentiation of bovine primary muscular cells and murine cells of the C2C12 lineage. We show that the transient surexpression of the murine Pofut1 enzyme delays the expression of myogenic transcription factors belonging to bHLH family: Myf5, MyoD, Myogenine and MRF4. These preliminary results open new and exciting perspectives of analysing the influence of *O*-fucosyltransferases during the complex process of myogenesis.

Key words : *O*-fucosylation, genomic organisation, transcript variants, phylogeny, *O*-fucosyltransferase activity, myogenesis, cattle